Buletin Ilmiah Math. Stat. dan Terapanya (Bimaster)
Volume 13, No. 1 (2024), hal 137-146

PEMODELAN TINGKAT PENGANGGURAN TERBUKA MENGGUNAKAN *GENERALIZED LINEAR MODEL*

Stepanus Armadi Mori, Shantika Martha, Siti Aprizkiyandari

INTISARI

Tingkat pengangguran terbuka menjadi aspek yang perlu diperhatikan di negara berkembang seperti Indonesia. Jawa Timur merupakan provinsi dengan penduduk nomor 2 tertinggi di Indonesia tahun 2022 tentnya tidak terlepas dengan permasalahan tingkat pengangguran terbuka. Tidak terserapnya tenaga kerja usia produktif serta kurangnya lapangan pekerjaan yang tersedia menyebabkan pengangguran di suatu daerah dapat menjadi meningkat. Oleh karena itu, perlu dilakukan sebuah analisis untuk mencegah meningkatnya tingkat pengangguran terbuka. Untuk dapat menjadi bahan evaluasi oleh pemerintah dilakukan pemodelan menggunakan generalized linear model menggunakan distribusi gamma untuk mengetahui faktor-faktor yang mempengaruhi tingkat pengangguran terbuka berdasarkan sektor ketenagakeraan, sektor sosial dan kependudukan, dan sektor pendidikan serta mengetahui faktor-faktor berdasarkan model terbaik yang memiliki pengaruh terhadap tingkat pengangguran terbuka di Jawa Timur tahun 2022. Penelitian ini menggunakan data di Jawa Timur dengan data sebanyak 38 kabupaten/kota dengan variabel yaitu tingkat pengangguran terbuka (Y), sektor ketenagakerjaan yakni tingkat partisipasi angkatan kerja (X1), pencari kerja terdaftar (X2), dan lowongan kerja terdaftar (X3), sektor sosial dan kependudukan yakni persentase penduduk miskin (X4), angka harapan hidup (X5), laju pertumbuhan penduduk (X6), dan kepadatan penduduk (X7), serta sektor pendidikan yakni rata-rata lama sekolah (X8), dan harapan lama sekolah (X9). Berdasarkan hasil penelitian ini, didapatkan sektor sosial dan kependudukan sebagai model terbaik berdasarkan nilai akaike's information criterian terkecil. Dari model tersebut dapat diketahui bahwa faktor-faktor yang memiliki pengaruh signifikan terhadap tingkat pengangguran terbuka adalah persentase penduduk miskin.

Kata Kunci : Metode Generalized Linear Model, Distribusi Gamma, Akaike's Information Criterian

PENDAHULUAN

Indonesia sebagai negara berkembang pasti tidak terlepas dari permasalahan seperti kesejahteraan dan pertumbuhan ekonomi. Permasalahan ini perlu ditindaklanjuti agar masyarakat yang tinggal di negara tersebut memiliki kesejahteraan hidup yang layak. Aspek penting yang perlu diperhatikan yaitu tingkat pengangguran terbuka yang terdapat di wilayah tersebut. Berdasarkan publikasi Badan Pusat Statistik (BPS), Jawa Timur adalah provinsi yang menduduki posisi ke-2 tertinggi tingkat pengangguran terbuka di Indonesia [1].

Pengangguran terbuka adalah orang yang telah masuk ke dalam usia angkatan kerja namun belum mendapatkan kesempatan untuk bekerja, sedang mencari pekerjaan, ataupun tidak bekerja [2]. Sumber daya manusia yang banyak namun tidak diseimbangkan dengan jumlah lapangan pekerjaan yang tersedia dapat menyebabkan kecilnya peluang seseorang mendapatkan pekerjaan. Selain itu, lapangan pekerjaan yang minim juga dapat mengakibatkan seseorang menjadi pengangguran. Jika hal ini terus terjadi, maka tenaga kerja yang masuk ke dalam usia produktif tidak dapat diserap dengan baik. Hal ini harusnya menjadi faktor utama yang harus diperhatikan pemerintah agar dapat mengatasi permasalahan yang berkelanjutan ini.

Sebuah analisis perlu dilakukan agar dapat dijadikan bahan acuan dalam menindaklanjuti permasalahan tingginya tingkat pengangguran di Jawa Timur. Perkembangan teori statistika telah mempengaruhi berbagai aspek kehidupan, contohnya permasalahan sosial dan kependudukan, ekonomi, kesehatan, dan masih banyak lagi. Hal ini dikarenakan statistika adalah disiplin ilmu yang digunakan sebagai alat untuk mengumpulkan, menyajikan, menganalisis data serta pengambilan keputusan.

Analisis regresi adalah metode yang dapat digunakan untuk menganalisis faktor-faktor yang memiliki pengaruh terhadap tingkat pengangguran terbuka. Analisis regresi merupakan satu diantara banyaknya materi yang terdapat pada statisika inferensial. Analisis regresi digunakan untuk membangun suatu model persamaan dan berfungsi untuk menganalisis hubungan antar variabel respon dan variabel prediktor yang memiliki pola hubungan dengan bentuk linear maupun tidak linear serta biasanya digunakan juga untuk memprediksi nilai dari variabel respon berdasarkan nilai-nilai variabel prediktor [3].

Metode *Generalized Linear Model* (GLMs) adalah metode yang digunakan ketika model regresi linear tidak dapat diterapkan saat model regresi tersebut tidak mampu mengatasi permasalahan pada saat variabel respon yang digunakan berupa data diskrit dan tidak berdistribusi normal. GLMs adalah bagian dari analisis regresi yang pada penerapannya bentuk kurva atau distribusi variabel responnya tidak harus berdistribusi normal namun masuk ke dalam distribusi keluarga eksponensial. Distribusi yang masuk ke dalam distribusi keluarga eksponensial adalah seperti distribusi gamma, distribusi poisson, distribusi binomial, serta distribusi binomial negativ [4].

Penelitian ini bertujuan untuk membentuk model menggunakan GLMs berdasarkan sektor ketenagakerjaan, sektor sosial dan kependudukan, serta sektor pendidikan menggunakan distribusi gamma. Berikutnya adalah mendapatkan model terbaik berdasarkan nilai *Akaike's Information Criterian* (AIC) terkecil sehingga diketahui faktor-faktor apa saja yang mempengaruhi tingat pengangguran terbuka di Jawa Timur tahun 2022. Penelitian sebelumnya dilakukan menggunakan data mengenai Covid-19 di Indonesia. Distribusi data pada penelitian ini menggunakan distribusi gaussian, distribusi poisson, dan distribusi gamma. Metode GLMs digunakan sebagai solusi dari regresi klasik yang tidak dapat digunakan karena data pada penelitian ini tidak memenuhi asumsi-asumsi pada regresi klasik. Penelitian ini menghasilkan model $\mathrm{y} = 23{,}06763 + 0{,}02082\ x$ yang dibentuk melalui metode *Generalized Linear Model* (GLMs) berdistribusi gaussian. Berdasarkan model yang terbentuk tersebut dapat ditarik kesimpulan bahwa jumlah kasus terkonfirmasi Covid-19 per hari di indonesia memiliki nilai yang signifikan terhadap jumlah pasien meninggal akibat Covid-19 per hari di Indonesia [5].

Langkah awal yang dilakukan dalam penelitian ini adalah melakukan identifikasi pada variabel respon. Jika variabel respon yang digunakan masuk ke dalam distribusi keluarga eksponensial yaitu distribusi gamma, maka dilanjutkan dengan pemilihan fungsi penghubung yang terdapat pada distribusi gamma yaitu *inverse*. Identifikasi yang dilakukan pada variabel respon menggunakan nilai p-value, jika p-value memiliki nilai lebih besar dari alpha atau p-value > 0,05 maka H_0 ditolak dengan kesimpulan bahwa data pada variabel respon berdistribusi gamma. Selanjutnya di lakukan pembentukan model GLMs menggunakan sebaran gamma yang pada penerapannya digunakan metode *Maximum Likelihood Estimation* (MLE) pada tiga sektor yang telah dipilih. Dari model yang telah didapatkan, kemudian dilakukan pengujian signifikansi terhadap model tersebut. Pengujian signifikansi yang dilakukan pertama adalah pengujian secara parsial yang digunakan untuk mengetahui apakah dari model terbaik yang didapatkan secara individu atau setiap variabel prediktornya memiliki nilai yang signifikan atau tidak terhadap variabel responnya. Selanjutnya jika dalam pemodelan masih terdapat model yang variabel prediktor yang memiliki nilai tidak signifikan, maka akan dilakukan pemodelan kembali menggunakan variabel prediktor yang signifikan. Setelah model didapatkan dan uji signifikansi parameter dilakukan maka dilanjutkan dengan pemilihan model terbaik menggunakan AIC terkecil dari model yang terbentuk. Setelah semua analisis dilakukan maka tahap terakhir yang dilakukan adalah melakukan interpretasi terhadap model terbaik yang didapatkan.

DISTRIBUSI KELUARGA EKSPONENSIAL

Dalam *Generalized Linear Model* (GLMs), variabel respon yang digunakan berpeluang untuk memiliki distribusi data lebih dari satu. Suatu variabel yang dikatakan masuk ke dalam keluarga eksponensial adalah jika memenuhi syarat bentuk persamaan berikut [6].

$$f(y|\theta,\phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\} \quad (1)$$

Untuk:

y : Variabel Respon

θ : Parameter Kanonik

ϕ : Parameter Dispersi

Parameter kanonik adalah parameter yang menggambarkan hubungan antara variabel prediktor dan variabel respon, sedangkan parameter dispersi adalah parameter tambahan yang berfungsi untuk mengatasi ketidaksepmurnaan sebaran data. Penghubung kanonik adalah fungsi penghubung yang didapatkan dari penurunan bentuk keluarga eksponensial.

METODE *GENERALIZED LINEAR MODEL*

Generalized Linear Model (GLMs) merupakan sebuah metode yang didapatkan dari perluasan model regresi linear yang pada penerapannya variabel prediktornya diasumsikan memeliki efek linear akan tetapi tidak diasumsikan memiliki distribusi tertentu. Dalam penggunaannya, GLMs memiliki keunggulan dibandingkan dengan model linear dimana bentuk kurva atau distribusi variabel responnya tidak diharuskan berdistribusi normal tetapi masuk ke dalam keluarga eksponensial. Adapun distribusi yang masuk ke dalam keluarga eksponensial adalah distribusi poisson, distribusi binomial, distribusi gaussian, distribusi gamma dan lain-lain [4].

GLMs yang merupakan transisi dari model linear akan dijabarkan melalui beberapa komponen berikut [7].

1. Komponen acak, dimana nilai-nilai pengamatan pada variabel respon diasumsikan saling bebas dari distribusi tertentu.
2. Komponen sistematik, dimana variabel prediktor dan parameter β diasumsikan linear dengan $\boldsymbol{\eta} = \boldsymbol{X\beta}$.
3. Fungsi penghubung, dimana berfungsi untuk menghubungkan variabel prediktor linear η dengan nilai harapan dari variabel respon μ yang dinyatakan sebagai $\eta_i = g(\mu_i) = \beta_0 + \sum_{j=1}^{n} \beta_j x_{ij} = \boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{T}} \boldsymbol{\beta}$.

Model GLMs yang terbentuk dapat dinyatakan pada Persamaan (2).

$$g(\mu_i) = \boldsymbol{x}_i^T \boldsymbol{\beta}, i = 1, 2, \ldots, m \quad (2)$$

Untuk:

$\boldsymbol{\beta}$: $\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_n \end{pmatrix}$ adalah vektor kolom berukuran $((n+1) \times 1)$

$\boldsymbol{x_i}$: $\boldsymbol{x_i} = \begin{pmatrix} 1 \\ x_{i1} \\ \cdots \\ x_{in} \end{pmatrix}$ adalah vektor kolon berukuran $((n+1) \times 1)$

n : adalah banyaknya variabel prediktor

m : adalah banyaknya amatan

Berdasarkan beberapa komponen yang telah dijabarkan, maka dapat diketahui bahwa dalam pembentukan model menggunakan GLMs perlu diperhatikaan beberapa hal yaitu, variabel prediktor yang digunakan bersifat linear, variabel responnya masuk ke dalam keluarga eksponensial, serta harus memiliki fungsi penghubung.

ESTIMASI PARAMETER PADA *GENERALIZED LINEAR MODEL*

Suatu variabel acak X dikatakan memiliki distribusi gamma jika fungsi kepadatan peluangnya adalah [8]:

$$f(x) = \begin{cases} \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)}, x > 0_{(\alpha,\beta>0)} \\ 0, x \le 0 \end{cases} \tag{3}$$

Untuk:

α : Parameter Skala

β : Parameter Bentuk

$\Gamma(\alpha)$: Fungsi Gamma

Parameter skala adalah parameter yang berfungsi untuk mengontrol sebaran data pada distribusi probabilitas, sedangkan parameter bentuk adalah parameter yang berfungsi untuk mengontrol bentuk distribusi probabilitas. Pada *Generalized Linear Model* (GLMs), metode yang digunakan dalam melakukan estimasi parameter adalah *Maximum Likelihood Estimation* (MLE). MLE adalah metode yang digunakan dengan cara memaksimumkan sebuah fungsi. Dimisalkan $x_1, x_2, \dots, x_n$ adalah sebuah sampel variabel acak, maka fungsi kepadatan peluangnya sebagai berikut:

$$f(x_1, x_2, \dots, x_n);\ \theta) = f(x_1;\theta)\, f(x_2;\theta) \dots f(x_i;\theta) \tag{4}$$

Langkah awal yang dilakukan adalah menentukan fungsi padat peluang peubah acak $(X_1, X_2, \dots, X_n)$ dari Persamaan (3). Langkah ini dilakukan dengan menggunakan Persamaan (3) dan Persamaan (4), sehingga:

$$\begin{aligned} f(x_i) &= f(\{x_1, x_2, \dots, x_n\}) \\ &= \prod_{i=1}^{n} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} e^{\frac{-\sum_{i=1}^{n} x_i}{\beta}} (x_1 x_2 \dots x_n)^{\alpha-1} \end{aligned} \tag{5}$$

Pada Persamaan (5), dibentuk fungsi kepadatan peluang pada Persamaan (5) ke dalam model $\mathrm{L}(x|\alpha, \beta)$ atau fungsi *likelihood* , sehingga:

$$\begin{aligned} L(x_i|\alpha,\beta) &= L(f(x_i|\alpha,\beta)) \\ &= \left(\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\right)^{n} e^{\frac{-\sum_{i=1}^{n} x_i}{\beta}} ((x_1 x_2 \dots x_n)^{\alpha-1})^{n} \end{aligned} \tag{6}$$

Selanjutnya dibentuk fungsi *likelihood* pada Persamaan (6) ke dalam model ln $\mathrm{L}(x|\alpha, \beta)$ atau memaksimumkan fungsi *likelihood*, sehingga:

$$\begin{aligned} L(x_i|\alpha,\beta) &= \ln L(f(x_i|\alpha,\beta)) \\ &= -n\ln\big(\Gamma(\alpha)\big) - n\,\alpha\ln(\beta) - \frac{\sum_{i=1}^{n} x_i}{\beta} + n\alpha\sum_{i=1}^{n}\ln(x_i) - n\sum_{i=1}^{n}\ln(x_i) \end{aligned} \tag{7}$$

Setelah didapatkan fungsi *likelihood* yang telah dimaksimumkan, maka akan dilakukan penurunan terhadap parameter α dan β serta menyamakannya dengan 0, sehingga:

$$\begin{aligned} \frac{\partial}{\partial\alpha}\ln L(x_i|\alpha,\beta) &= 0 \\ \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} + \ln(\hat{\beta}) - n\sum_{i=1}^{n} &= 0 \end{aligned} \tag{8}$$

dan untuk β.

$$\frac{\partial}{\partial\beta}\ln L(x_i|\alpha,\beta) = 0$$

$$-n\hat{\alpha}\frac{1}{\hat{\beta}} + \left(\sum_{i=1}^{n}(x_i)\right)\frac{1}{\hat{\beta}^2} = 0 \tag{9}$$

Selanjutnya Persamaan (8) dan Persamaan (9) dapat ditulis sebagai berikut:
Persamaan (8) menjadi

$$\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} + \ln\hat{\beta} = n\sum_{i=1}^{n}\ln(x_i) \tag{10}$$

Persamaan (9) menjadi

$$\frac{\sum_{i=1}^{n} x_i}{\hat{\beta}^2} = \frac{n\hat{\alpha}}{\hat{\beta}}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i}{n\hat{\alpha}} \tag{11}$$

Persamaan (10) dan Persamaan (11) selanjutnya disubstitusikan sehingga dapat dinyatakan pada Persamaan (12)

$$\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} + \ln\frac{\sum_{i=1}^{n} x_i}{n\hat{\alpha}} = n\sum_{i=1}^{n}\ln(x_i)$$

$$\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} + \ln\sum_{i=1}^{n} x_i = n\sum_{i=1}^{n}\ln(x_i) - \ln n\hat{\alpha} \tag{12}$$

UJI PARSIAL

Pada analisis regresi perlu lakukan pengujian untuk mengetahui pengaruh antar variabel. Adapun variabel yang akan diuji adalah variabel prediktor terhadap variabel respon. Uji parsial dilakukan untuk melihat hubungan yang dilakukan secara individu atau setiap variabel prediktor terhadap variabel respon [9]. Adapun hipotesis yang digunakan sebagai berikut:

H_0 : $\beta_j = 0$ (variabel prediktor tidak memiliki pengaruh yang signifikan terhadap variabel respon)

H_1 : $\beta_j \neq 0$, dengan $j = 1, 2, \ldots, k$ (variabel prediktor memiliki pengaruh yang signifikan terhadap variabel respon

Statistik uji yang digunakan adalah:

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{13}$$

dengan tingkat signifikansi sebesar α, maka kriteria keputusannya tolak H_0 jika $|t| \geq t_{\left(n-p-1;\frac{\alpha}{2}\right)}$.

UJI SIMULTAN

Uji simultan atau serentak adalah pengujian yang dilakukan dalam melakukan pemeriksaan signifikansi pada sebuah parameter. Pemeriksaan signifikansi yang dimaksudkan terjadi pada parameter β terhadap variabel respon secara bersamaan. Pada *generalized linear model*, digunakan uji *likelihood ratio test* dengan hipotesis yang digunakan adalah sebagai berikut:

H_0 : $\beta_j = 0$ (tidak terdapat satupun parameter β yang memiliki pengaruh signifikan terhadap variabel respon)

H_1 : $\beta_j \neq 0$, dengan $j = 1, 2, \ldots, k$ (setidaknya terdapat satu parameter β yang memiliki pengaruh signifikan terhadap variabel respon

Statistik uji yang digunakan adalah [10]:

$$G = -2\left[\frac{L_0}{L_1}\right] \sim X^2_{(\alpha,v)} \tag{14}$$

L_0 : Fungsi *likelihood* dari model tanpa melibatkan variabel prediktor

L_1 : Fungsi *likelihood* dari model yang melibatkan variabel prediktor dengan $\alpha = 0{,}05$ dan v adalah jumlah variabel prediktor kriteria pengujian adalah tolak H_0 jika nilai $G > X^2_{(\alpha,v)}$.

AKAIKE'S INFORMATION CRITERIAN

Dalam GLMs, pemilihan model terbaik yang khalayak digunakan adalah *Akaike's Infoemation Criterian* (AIC). AIC merupakan sebuah kriteria yang *goodnees of fit* modelnya diseimbangkan. *Goodnees of fit* yang diseimbangkan berdasarkan nilai *likelihood* dengan banyaknya parameter. Pada penggunaannya,pemilihan model pada AIC di ambil berdasarkan nilai yang terkecil, selain itu biasanya AIC digunakan berdasarkan perbandingan dua atau lebih model yang terbentuk. Nilai AIC dapat dicari menggunakan persamaan berikut [3].

$$AIC = -2l(\theta) + 2q \tag{16}$$

dengan $l(\theta)$ adalah nilai likelihood yang akan dibentuk dan q adalah banyaknya parameter dalam model.

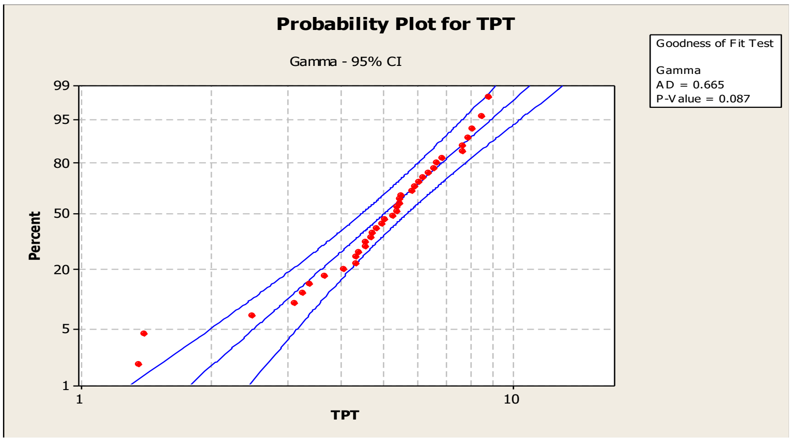
HASIL DAN PEMBAHASAN

Pada penelitian ini, digunakan data sekunder melalui publikasi Badan Pusat Statistik Jawa Timur. Data yang digunakan adalah data sebanyak 38 kabupaten/kota pada tahun 2022 dengan tingkat pengangguran terbuka (Y) sebagai variabel respon, tingkat partisipasi angkatan kerja (X1), pencari kerja terdaftar (X2), lowongan kerja terdaftar (X3), persentase penduduk miskin (X4), angka harapan hidup (X5), laju pertumbuhan penduduk (X6), kepadatan penduduk (X7), rata-rata lama sekolah (X8), dan harapan lama sekolah (X9) sebagai variabel prediktor. Berikut merupakan statistik deskriptif yang diperoleh bersadarkan masing-masing variabel.

Tabel 1. Statistik Deskriptif Variabel Respon dan Variabel Prediktor

Variabel Prediktor	Mean	Minimum	Maximum
Y	5,27	1,36	8,80
X1	71,28	63,08	82,99
X2	3055,45	563	12121
X3	8592,24	413	60017
X4	10,31	3,79	21,61
X5	72,08	67,29	74,54
X6	0,73	0,16	1,41
X7	2017,13	414	8596
X8	8,27	5,06	11,67
X9	13,48	11,91	15,76

Berdasarkan Tabel 1 dapat diketahui bahwa tingkat pengangguran terbuka memiliki nilai terendah sebesar 1,36 pada kabupaten Sumenep dan nilai tertinggi sebesar 8,80 pada kabupaten Siduarjo. Tingkat partisipasi angkatan kerja memiliki nilai terendah sebesar 63,08 pada kota Malang dan nilai tertinggi pada kabupaten Pacitan. Pencari kerja terdaftar memiliki nilai terendah sebesar 563 pada kota Mojokerto dan nilai tertinggi sebesar 12121 pada kota Surabaya. Lowongan kerja terdaftar memiliki nilai terendah sebesar 413 pada kabupaten Tulungagung dan nilai tertinggi sebesar 60017 pada kota Surabaya. Lowongan kerja terdaftar memiliki nilai terkecil sebesar 3,79. Persentase penduduk miskin memiliki nilai terendah sebesar 3,79 pada Kota Batu dan nilai tertinggi sebesar 21,61 pada Kabupaten Sampang. Angka harapan hidup memiliki nilai terendah sebesar 67,29 pada kabupaten Bondowoso dan nilai tertinggi sebesar 74,54 pada kabupaten Tulungagung. Laju pertumbuhan penduduk memiliki nilai terkecil sebesar 0,16 pada kota Malang dan nilai tertinggi sebesar 1,41 pada kabupaten Bangkalan. Kepadatan penduduk memiliki nilai terendah sebesar 414 pada kabupaten Pacitan dan nilai tertinggi sebesar 8596 pada kota Surabaya. Rata-rata lama sekolah memiliki nilai terendah sebesar 5,06 pada kabupaten Sampang dan nilai tertinggi sebesar 11,67 pada kota Madiun. Serta harapan lama sekolah memiliki nilai terendah sebesar 11,91 pada kabupaten Bangkalan dan nilai tertinggi sebesar 15,76 pada kota Malang.



Gambar 1. Uji Identitas Distribusi

Berdasarkan Gambar 1 dapat diketahui bahwa hasil pengujian identifikasi distribusi pada variabel respon menunjukkan nilai p-value 0.087 > 0.05 yang berarti tolak H_0 dengan kesimpulan data yang digunakan pada variabel respon berdistribusi gamma.

Tabel 2. Rangkuman Uji Signifikansi Koefisien dan Konstanta Untuk 3 Model

Model	Parameter	Koefisien	Std. Error	t-value	Pr(>|t|)
1	(Intercept)	-3,469e-01	2,080e-01	-1,668	0,105
	X1	7,775e-03	2,956e-03	2,630	0,013*
	X2	2,256e-06	6,158e-06	0,366	0,716
	X3	-2,008e-06	1,151e-06	-1.744	0,090
2	(Intercept)	0,956772	0,458848	2,085	0,045*
	X4	0,007737	0,002716	2,848	0,007*
	X5	-0,009899	0,006140	-1612	0,116
	X6	-0,114778	0,037049	-3,098	0,004*
	X7	-0,013827	0,004940	-2,799	0,008*
3	(Intercept)	0,229229	0,175569	1,306	0,200
	X8	-0,027209	0,009574	-2,842	0,007*
	X9	0,014237	0,017171	0,829	0,413

***Sig 0,05**

Berdasarkan Tabel 2, dapat diketahui bahwa untuk masing-masing sektor masih terdapat beberapa variabel prediktor yang memiliki nilai tidak signifikan. Oleh sebab itu, perlu dilakukan pemodelan kembali dengan hanya menggunakan variabel prediktor yang memiiki nilai signifikan sehingga didapatkan hasil pengujian seperti pada Tabel 3 berikut.

Tabel 3. Rangkuman Uji Signifikansi Koefisien dan Konstanta Untuk 3 Model Terbaik

Model	Parameter	Koefisien	Std. Error	t-value	Pr(>|t|)
1	(Intercept)	-0.4071	0.2140	-1.902	0.0651
	X1	0.0084	0.0030	2.772	0.0087*
2	(Intercept)	0.1001	0.0260	3.853	0.0004*
	X4	0.0093	0.0027	3.456	0.0014*
3	(Intercept)	0.3680	0.0516	7.133	2.2e-08*
	X8	-0.0208	0.0057	-3.644	0.0008*

*Sig. 0,05

Dari Tabel 3, model terbaik yang terbentuk berdasarkan masing-masing sektor adalah sebagai berikut:

Model 1 atau sektor ketenagakerjaan

$$\hat{\mu} = ((-0{,}4071) + 0{,}0084\ X_1)^{-1}$$

Model 2 atau sektor sosial dan kependudukan

$$\hat{\mu} = (0{,}1001 + 0{,}0093\ X_4)^{-1}$$

Model 3 atau sektor pendidikan

$$\hat{\mu} = (0{,}3680 - 0{,}0208\ X_8)^{-1}$$

Berdasarkan Tabel 3, dapat diketahui bahwa pengujian secara parsial pada sektor ketenagakerjaan yaitu tingkat partisipasi angkatan kerja memiliki pengaruh yang signifikan secara positif terhadap tingkat pengangguran terbuka di provinsi Jawa Timur. Pada sektor sisoal dan kependudukan yaitu persentase penduduk miskin memiliki pengaruh yang signifikan secara positif terhadap tingkat pengangguran terbuka di jawa timur. Pada sektor pendidikan yaitu rata-rata lama sekolah memiliki pengaruh yang signifikan secara negative terhadap tingkat pengangguran terbuka di provinsi Jawa Timur.

Setelah dilakukannya pemodelan dan didapatkan model terbaik untuk masing-masing sektor, maka dilakukan pemilihan model terbaik *generalized linear model* menggunakan distribusi gamma dari ketiga model tersebut berdasarkan nilai AIC terkecil. Nilai AIC yang terbentuk dapat dilihat pada Tabel 4.

Tabel 4. Nilai AIC dari Model

Model	AIC
Model 1.1	155,04
Model 2.1	149,97
Model 3.1	151,35

Berdasarkan Tabel 4 dapat diketahui bahwa nilai AIC untuk Model 1 sebesar 155,04, nilai AIC untuk Model 2 sebesar 149,97, dan nilai AIC untuk Model 3 sebesar 151,35. Selain itu, dapat diketahui jugan bahwa nilai AIC terkecil dimiliki oleh model 2 dengan nilai sebesar 149,97. Model 2 merupakan pemodelan yang terbentuk dari pemodelan yang dilakukan pada sektor sosial dan kependudukan. Sehingga pada penelitian ini dapat ditarik kesimpulan bahwa model terbaik dari masing-masing model terbaik yang terbentuk pada pemodelan tingkat pengangguran terbuka menggunakan *generalized linear model* dengan distribusi gamma adalah Model 2 yang merupakan pemodelan pada sektor sosial dan kependudukan.

PENUTUP

Berdasarkan penelitian yang telah dilakukan, didapatkan tiga model yaitu model dengan sektor ketenagakerjaan, model dengan sektor sosial dan kependudukan, serta model dengan sektor pendidikan. Dari masing-masing model tersebut didapatkan model terbaik berdasarkan nilai AIC terkecil yaitu model dengan sektor sosial dan ketenaga kerjaan dengan persentase penduduk miskin sebagai faktor yang memiliki pengaruh secara signifikan terhadap tingkat pengangguran terbuka di Jawa Timur tahun 2022. Selain itu, dapat diketahui bahwa jika variabel lain memiliki nilai konstan, maka tingkat pengangguran terbuka akan berubah dengan sendirinya sebesar nilai konstanta 0,1001. Serta jika variabel lain bernilai konstan, maka tingkat pengangguran terbuka akan bertambah sebesar 0,0093 untuk setiap satu satuan persentase penduduk miskin.

DAFTAR PUSTAKA

[1]. Badan Pusat Statistik, 2023, *Indikator*, BPS Provinsi Jawa Timur.

[2]. Badan Pusat Statistik Jawa Timur, 2023, Provinsi Jawa Timur Dalam Angka 2023, BPS Provinsi Jawa Timur

[3]. Kusnandar, D., 2004, *Metode Statistik dan Aplikasinya dengan Minitab dan Excel*, Madyan Press, Yogyakarta.

[4]. Jamilatuzzahro, Caraka, R. E., Herliansyah. R., 2018, *Aplikasi Generalized Linear Model pada R*, Ed.1, Innosain, Yogyakarta.

[5]. Saidi, S., Herawati, N., dan Nisa, K., Modeling with generalized linear model on covid-19: cases in Indonesia, *International Journal of Electronics and Communications System*, 1(1): 25-32.

[6]. De Jong, P., and Heller, Z.G., 2008. *Generalized linear models for insurance data*, Cambridge: Cambridge University Press.

[7]. McCullough, P., & Nelder, J. A., 1989, *Generalized Linear Models Chapman And Hall*. New York.

[8]. Hogg., McKean., dan Craig., 2005. *Introduction to Mathematical Statistic*. Sixth edition. New Jersey: Pearson Prentice Hall.

[9]. Kusnandar, D., Debataraja, N. N., Mara, M. N., dan Satyahadewi. N., 2019. *Metode Statistika serta Aplikasinya dengan Minitab, Excel dan R*. Untan Press, Pontianak.

STEPANUS ARMADI MORI	:	Jurusan Matematika, FMIPA UNTAN, Pontianak, samori.ss45@student.untan.ac.id
SHANTIKA MARTHA	:	Jurusan Matematika, FMIPA UNTAN, Pontianak, Shantika.martha@math.untan.ac.id
SITI APRIZKIYANDARI	:	Jurusan Budidaya Pertanian, FAPERTA UNTAN, Pontianak, siti.aprizkiyandari @faperta.ac.id