
PEMODELAN KEPADATAN PENDUDUK DENGAN PENDEKATAN REGRESI POISSON TERGENERALISASI

Alfin Dratama Maulana^{1*}, Dewi Anggraini², Yuana Sukmawaty³

^{1,2,3} Program Studi Statistika Fakultas MIPA, Universitas Lambung Mangkurat, Indonesia

*e-mail alfinmaulana31@gmail.com

Abstract

Population density tends to increase and land availability decreases every year due to migration. Efforts to improve the quality of the population will be more difficult to do in areas with high densities, so that it can cause problems both in the socio-economic field, land availability, and food needs. Therefore, it is necessary to know the factors that affect population density. One approach to predicting the relationship pattern between factors and population density, can use Poisson regression. Poisson regression is one of the nonlinear regressions that is often used to model variables in the form of whole numbers. This study aims to describe the characteristics of the population of each province in Indonesia, estimate the parameters of the Poisson regression model and explain the factors that have a significant influence on the problem of population density. Factors that are thought to have an effect on population density are population size, area size, number of elementary school level, number of junior high school level school, number of high school level school, number of puskesmas, number of cooperatives and number of deaths. The results of this study indicate that the generalized Poisson regression model $\hat{\mu}_i(X_i) = e^{(8.032+0.000097X_1-0.000136X_2)}$, that the factors that influence the population density of each province in Indonesia at the level of population density (Y) and population (X_1) has a unidirectional relationship, meaning that for every increase of one population, the population density increases by 9.7%, while the area (X_2) has an opposite relationship of 0.0136%, which means that one unit is narrower area, then the level of population density increases by 0.0136%.

Keywords: Poisson regression, population density, correlation coefficient

1. PENDAHULUAN

Penduduk adalah kumpulan manusia yang menempati wilayah tertentu yang dapat berubah sewaktu-waktu karena adanya proses kelahiran, kematian, dan perpindahan dari satu wilayah ke wilayah lain. Apabila kelahiran penduduk semakin tinggi pada suatu desa atau wilayah tanpa diimbangi penyebaran penduduk akan menyebabkan kepadatan penduduk (Dewi, 2021). Menurut Badan Pusat Statistik (2010), kepadatan penduduk adalah penduduk yang mendiami wilayah suatu daerah persatuan luas, kepadatan penduduk merupakan persoalan penting bagi suatu negara karena dampaknya sangat besar terhadap suatu kemajuan di negara tersebut (Pakpahan, 2018).

Hal ini memicu suatu ledakan penduduk di daerah-daerah tertentu terutama di daerah yang mempunyai daya tarik cukup kuat contohnya baik dari segi ekonomi, fasilitas

sosial yang memadai, jaminan keamanan, maupun dari aspek sosial. Hal ini menjadi masalah yang sangat lazim bagi kehidupan manusia karena mempunyai kecenderungan mencari tempat-tempat yang dekat dengan sumber penghidupannya seperti dekat dengan industri, dekat jalan raya, supermarket dan lain sebagainya (Suhardi, 2009). Berdasarkan publikasi yang ada di dukcapil.kemendagri.go.id penduduk Indonesia terus mengalami perkembangan yang sangat signifikan berdasarkan data administrasi kependudukan (Adminduk) per 2021, jumlah penduduk Indonesia adalah sebanyak 272.229.372 jiwa, dimana 137.521.557 jiwa adalah laki-laki dan 134.707.815 jiwa adalah perempuan. Dari total 272 juta tersebut ada sebesar 56,01 persen terpusat di pulau Jawa, dimana Provinsi Jawa Barat (Jabar) merupakan Provinsi dengan jumlah penduduk terbanyak se Indonesia sebanyak 47.586.943 jiwa. Sedangkan Provinsi dengan penduduk tersedikit adalah Kalimantan Utara (Kaltara) sebanyak 692.239 jiwa. Oleh karena itu adapun kabupaten yang penduduknya terbanyak di Indonesia adalah Kabupaten Bogor sebanyak 5.198.693 jiwa.

Sementara kabupaten dengan penduduk tersedikit adalah Kabupaten Supiori di Papua sebanyak 24.719 jiwa. Hal ini tentu saja menjadi permasalahan di Indonesia karena pertambahan penduduk tidak diikuti dengan meningkatnya sektor ekonomi dan fasilitas sosial yang memadai. Berdasarkan maraknya kepadatan penduduk dan ketersediaan lahan yang semakin berkurang setiap tahunnya akibat migrasi penduduk maka dari itu perlu sekiranya untuk mengetahui apa saja yang mempengaruhi kepadatan penduduk.

2. TINJAUAN PUSTAKA

2.1. Regresi Poisson

Regresi Poisson adalah salah satu regresi nonlinier yang sering digunakan untuk menampilkan hubungan antara variabel dependen berupa data diskrit dan variabel independen yang berupa data kontinu. Regresi Poisson adalah penerapan Generalized Linear Model (GLM) yang menggambarkan hubungan antara variabel dependen Y sebagai informasi diskrit atau menghitung informasi dan faktor otonom sebagai informasi diskrit, kontinu, benar atau campuran (Ruliana, 2016).

2.2. Pendugaan Parameter Model Regresi Poisson

Strategi Maximum likelihood estimation (MLE) adalah teknik untuk mengetahui nilai batas mana yang memperluas kemampuan probabilitas. Rata-rata dalam Regresi Poisson ditampilkan sebagai komponen dari berbagai faktor bebas. Pertama-tama, penting untuk memutuskan kemampuan probabilitas dan kondisi probabilitas yang dapat digunakan untuk mengukur parameter dalam regresi Poisson (Nugraha, 2013). Fungsi likelihood untuk distribusi Poisson adalah:

$$L(y, \mu) = \prod_{i=1}^n f(y_i, \lambda) \quad (1)$$

$$L(y, \lambda) = \prod_{i=1}^n \left\{ \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \right\} = \frac{\{\prod_{i=1}^n \lambda^{\sum_{i=1}^n y_i}\}}{\prod_{i=1}^n y_i!} \quad (2)$$

Penilaian parameter menggunakan pendekatan model linier dengan strategi MLE adalah turunan utama dari kapabilitas probabilitas log, sehingga struktur yang akan digunakan untuk mendapatkan parameter menggunakan teknik MLE sebagai berikut:

$$\frac{\partial 1nL(y, \beta)}{\partial \beta} = 0 \quad (3)$$

Sehingga dipunyai

$$L(y, \beta) = L \sum_{k=1}^n y_i \ln \hat{\beta}_0 + \sum_{j=1}^m \beta_j (x_{ij}) - \sum_{k=1}^n (\hat{\beta}_0 + \sum_{j=1}^m \beta_j X_{ij}) - \sum_{k=1}^n 1 (y_k!) \quad (4)$$

Turunan kedua

$$\frac{\partial^2 L(L_y, \beta)}{\partial^2 \beta_k} = - \sum_{i=1}^n \frac{x_{ik}^2 y_i}{(\sum_{j=1}^m x_{ij} \beta_j)^2} \quad (5)$$

$$\frac{\partial^2 L(L_y, \beta)}{\partial \beta_1 \partial \beta_k} = - \sum_{i=1}^n \frac{x_{ik} x_i y_i}{(\sum_{j=1}^m x_{ij} \beta_j)^2} \quad (6)$$

$$\frac{\partial L(y, \beta)}{\partial^2 \beta_0} L = - \sum_{i=1}^n \frac{y_i}{(\beta_0 + \beta_1 X_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i})^2} \quad (7)$$

2.3. Uji Statistik G

Uji statistik G adalah pengujian parameter secara simultan model regresi Poisson digunakan untuk melihat ada tidaknya pengaruh variabel independen terhadap variabel dependen (Darnah, 2011). Pengujian parameter model regresi Poisson dilakukan dengan menggunakan metode Maximum Likelihood Ratio Test (MLRT). Dengan statistik uji:

$$G - 2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 [\ln L(\hat{\Omega}) - \ln L(\hat{\omega})] \quad (8)$$

Hipotesis:

$$H_0 : \beta_1 = \beta_2 = \beta_3 \dots \beta_8 = 0$$

H_1 : seringkali ada satu β_j yang berpengaruh terhadap model

2.3 Uji Signifikansi Parameter Individual (Uji Wald)

Uji wald merupakan suatu pengujian secara parsial yang bertujuan untuk mengetahui apakah pengaruh yang diberikan masing-masing variabel independen digunakan pengujian secara parsial (Darnah, 2011). Statistik uji yang digunakan adalah uji wald berikut adalah statistik ujinya:

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (9)$$

$H_0 : \beta_i = 0$ (tidak ada pengaruh variabel X terhadap variabel Y)

$H_1 : \beta_i \neq 0$ (ada pengaruh variabel X terhadap variabel Y)

2.4 Uji Multikolinearitas

Tujuan digunakannya uji multikolinearitas dalam penelitian ini untuk mengetahui arah dan kekuatan hubungan antar variabel independen dan dependen. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel independen (Susanti, Sri dkk.2018) Hipotesis:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (10)$$

H_0 : Tidak ada korelasi antar variabel r_s dan r_{tabel}

H_1 : Adanya korelasi antar variabel r_s dan r_{tabel}

2.5 Equidispersi

Equidispersi adalah asumsi yang harus dipenuhi dalam regresi poisson, dimana nilai rata-rata dari varians pada variabel dependen harus bernilai sama. Menurut Harahap, R. I. (2020) pada pengujian *equidispersi* kadang terjadi pelanggaran asumsi dalam analisis data. Berikut adalah statistik ujinya:

$$Var(Y) = E(Y) \quad (11)$$

Hipotesis:

$H_0 : Var(Y)$ besar dari $E(Y)$

$H_1 : Var(Y)$ bernilai sama $E(Y)$

Jika $Var(Y) > E(Y)$ maka terima H_0 variabel dependen tidak terdapat *equidispersi*

2.6 Overdispersi dan Underdispersi

Dalam model Regresi Poisson ada beberapa praduga yang harus dipenuhi, salah satunya adalah kemiripan antara rata-rata variabel dependen dengan nilai faktor bebas dan perubahannya, atau disebut *equidispersion*. Dalam pemeriksaan data hitungan, seringkali diamati data yang variansnya lebih besar dari rata-ratanya. masalah seperti ini disebut dengan *overdispersi*. ada dua cara untuk mendeteksi *overdispersion* yaitu:

Nilai Deviance

$$\theta_1 = \frac{D}{db} > 1; D = 2 \sum_{i=1}^n y_i \log \left\{ \frac{y_i}{\mu_i} \right\} \quad (12)$$

Nilai Chi-Square

$$\theta_2 = \frac{x^2}{db} > 1; \chi^2 = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\sigma_i} \quad (13)$$

Jika θ_1 dan θ_2 bernilai lebih dari 1 terjadi *Overdispersi*.

2.7 Regresi Poisson Tergeneralisasi

Regresi Poisson Tergeneralisasi adalah suatu model regresi yang digunakan untuk menganalisis hubungan antara sebuah variabel dependen yang berupa data cacah dengan satu atau lebih variabel independen. Regresi Poisson Tergeneralisasi juga dapat digunakan baik dalam keadaan *underdispersi*, *equidispersi* ataupun *Overdispersi* (Fadhillah, 2011). Adapun fungsi peluangnya adalah

$$f(y_i; \mu, a) = \left(\frac{\mu_i}{1 + \alpha \mu_i} \right)^{y_i} \frac{(1 + \alpha y_i)^{y_i - 1}}{y_i!} \exp \left(\frac{\mu_i (1 + \alpha y_i)}{1 + \alpha \mu_i} \right) \quad (14)$$

2.8 Akaike Information Criteria (AIC)

Metode Akaike Information Criteria atau AIC adalah metode yang dapat digunakan untuk memilih model regresi terbaik. Namun, pada penelitian ini ukuran AIC dipakai untuk melihat nilai kesalahan dalam membangun model. Menurut (Ruliana, 2016), AIC adalah sarana yang mampu menunjukkan seberapa tepat model tersebut dengan data yang dimiliki secara mutlak. Berikut adalah statistik ujinya:

$$AIC = -2 \ln L(\hat{\theta}) + 2k \quad (15)$$

3. METODE PENELITIAN

Penelitian ini data yang digunakan adalah data sekunder yang diperoleh dari website Badan Pusat Statistik <https://www.bps.go.id/>. Unit observasi yang digunakan dalam penelitian ini adalah 34 provinsi di Indonesia pada tahun 2019. Variabel yang digunakan dalam penelitian ini terdiri atas satu variabel dependen (Y) sedangkan variabel independen (X) terdiri dari 8 variabel

Tabel 1 Variabel Penelitian

Nama Variabel	Label	Tipe Variabel
Y	Kepadatan Penduduk	Diskrit
X₁	Jumlah Penduduk	Diskrit
X₂	Luas Wilayah	kontinu
X₃	Jumlah Sekolah Tingkat SD	Diskrit
X₄	Jumlah Sekolah Tingkat SMP	Diskrit

X₅	Jumlah Sekolah Tingkat SMA	Diskrit
X₆	Jumlah Puskesmas	Diskrit
X₇	Jumlah Koperasi	Diskrit
X₈	Jumlah Kematian	Diskrit

Langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

- 1) Mendeskripsikan permasalahan kepadatan penduduk di Indonesia dengan data tahun 2019 menggunakan ukuran pemusatan data (rata-rata) dan ukuran penyebaran data (minimal, maksimal, std deviasi).
- 2) Melakukan uji korelasi spearman's rho untuk menguji keeratan hubungan antar variabel independen dan menguji bentuk sebaran variabel dependen.
- 3) Melakukan pengujian *equidispersi*, jika asumsi distribusi Poisson tidak terpenuhi maka dilanjutkan dengan pengujian *Overdispersi*
- 4) Selanjutnya estimasi parameter model regresi Poisson tergeneralisasi.
- 5) Menguji Kesignifikansi parameter model, menggunakan Maximum Likelihood Estimation
- 6) Menentukan Model Generalized Regresi Poisson
- 7) Melakukan inferensi model regresi Poisson dengan uji serentak (Uji G) dan uji parsial (uji Wald), dan menghitung nilai AIC pada model tersebut.
- 8) Pembahasan dan interpretasi hasil.

4. HASIL DAN PEMBAHASAN

4.1 Karakteristik Permasalahan Kepadatan Penduduk di Indonesia

Pada penelitian ini deskripsi statistik dari variabel penelitian yang digunakan disajikan pada Tabel 4.1 sebagai berikut:

Tabel 2 Statistika Deskriptif dari Variabel Penelitian

Variabel	Rata - rata	Minimum	Maksimum	Standar Deviasi
Kepadatan Penduduk(Y) Jiwa/km ²	713	9	15900	2708,77
Jumlah Penduduk (X ₁) Ribuan Jiwa	7.850	696	49.023	11166,16
Luas Wilayah (X ₂) km ²	56379,61	664.01	319036.05	61062,75
Jumlah Sekolah Tingkat SD (X ₃) (Unit)	4.395	478	19718	5068,89
Jumlah Sekolah Tingkat SMP (X ₄)	1.193	180	5448	1202.98
Jumlah Sekolah Tingkat SMA (X ₅) (Unit)	410	61	1663	380.95

Jumlah Puskesmas (X_6) (Unit)	179	28	781	155.79
Jumlah Koperasi (X_7) (Unit)	3618	476	21757	4278,77
Jumlah Kematian (X_8) (Unit)	563	123	2450	555,81

Sumber: SPSS 24

Berdasarkan dari hasil Tabel 4.1 dapat diketahui bahwa di 34 provinsi untuk Kepadatan Penduduk (Y) pada tahun 2019 memiliki rata-rata sebesar 713 Jiwa/km². Provinsi DKI Jakarta merupakan provinsi dengan Kepadatan Penduduk (Y) tertinggi dengan nilai 15.900 Jiwa/km², sementara itu provinsi papua barat merupakan provinsi dengan kepadatan penduduk terendah dengan 9 Jiwa/km² dari kedua provinsi tersebut DKI Jakarta merupakan provinsi yang memiliki jumlah kepadatan penduduk dengan nilai tertinggi di atas rata-rata. Sedangkan pada Tabel 4.1 variabel Jumlah Penduduk (X_1) provinsi Jawa Barat merupakan provinsi yang memiliki jumlah penduduk tertinggi di tahun 2019 dengan nilai 49.023 Ribu Jiwa sementara itu provinsi kalimantan utara merupakan provinsi dengan jumlah penduduk terendah dengan 696 Ribu Jiwa. Pada Tabel 4.1 variabel Luas Wilayah (X_2) dapat diketahui provinsi yang memiliki luas wilayah terluas yaitu provinsi Papua dengan luas 319.036,05 km² sedangkan provinsi DKI Jakarta memiliki luas wilayah terkecil dengan nilai 664.01 km².

Berdasarkan Tabel 4.1 Variabel X_3 , X_4 , X_5 memiliki nilai tertinggi dan terendah pada provinsi yang sama. Jumlah Sekolah tingkat SD (X_3) pada provinsi Jawa Barat merupakan daerah dengan nilai tertinggi, sebesar 19.718 Unit dan pada provinsi Kalimantan Utara merupakan daerah dengan nilai terendah sebesar 478 Unit. Jumlah Sekolah tingkat SMP (X_4) pada provinsi Jawa Barat merupakan daerah dengan nilai tertinggi sebesar 5.448 Unit dan pada provinsi Kalimantan Utara merupakan daerah dengan nilai terendah sebesar 180 Unit. Jumlah Sekolah tingkat SMA (X_5) pada provinsi Jawa Barat merupakan daerah dengan nilai tertinggi sebesar 1.663 Unit dan pada provinsi Kalimantan Utara merupakan daerah dengan nilai terendah sebesar 61 Unit. Sehingga dapat disimpulkan bahwa provinsi Jawa Barat merupakan provinsi yang memiliki jumlah sekolah dengan nilai tertinggi di atas rata-rata.

Variabel Jumlah Puskesmas (X_6) pada provinsi Jawa Barat merupakan daerah dengan nilai tertinggi, sebesar 781 Unit dan pada provinsi Kalimantan Utara merupakan daerah dengan nilai terendah sebesar 28 Unit. Jumlah Koperasi Aktif (X_7) pada provinsi Jawa Timur merupakan daerah dengan nilai tertinggi, sebesar 21.757 Unit dan pada provinsi Kalimantan Utara merupakan daerah dengan nilai terendah sebesar 476 Unit. Jumlah Kematian (X_8) pada provinsi Jawa Tengah merupakan daerah dengan nilai tertinggi, sebesar 2.450 Jiwa dan pada provinsi Papua barat merupakan daerah dengan nilai terendah sebesar 123 Jiwa.

4.2 Uji Korelasi Spearman'rho Antar Variabel Independen

Uji korelasi antara variabel penelitian independen (X) dengan variabel independen(X) yang lain (multikolinieritas) dilakukan menggunakan uji korelasi Rank Spearman Rho. Sebagai ilustrasi diambil 2 (dua) buah variabel independen yaitu X_5 dan X_6 . Koefisien korelasi antara kedua variabel tersebut menghasilkan nilai sebagai berikut:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$r_s = 1 - \frac{6(1105)}{34(34^2 - 1)}$$

$$r_s = \frac{66}{77}$$

$$r_s = 0,831$$

H_0 : Tidak ada korelasi antar variabel X_5 dan X_6

H_1 : Adanya korelasi antar variabel X_5 dan X_6

Jika $r_s \geq r_{tabel} = 0,339$, maka H_0 ditolak. dan terima H_1 maka terdapat multikolinearitas.

Dapat dilihat pada hasil persamaan 20, nilai $r_s = 0,831 > r_{tabel} = 0,339$ sehingga H_0 ditolak. dan terima H_1 maka terdapat multikolinearitas antara Jumlah Sekolah Tingkat SMA (X_5) dan Jumlah Puskesmas (X_6).

Dengan cara yang sama akan diperoleh koefisien korelasi antar variabel penelitian antar variabel independen lainnya sebagaimana tercantum pada Tabel 4.2

Tabel 3. Nilai korelasi Sperman'rho untuk variabel independen

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	1.000	0.018	0.921**	0.891**	0.902**	0.723**	0.808**	0.716**
X_2	0.018	1.000	0.242	0.217	0.220	0.251	0.065	0.288
X_3	0.921**	0.242	1.000	0.952**	0.919**	0.756**	0.765**	0.830**
X_4	0.891**	0.217	0.952**	1.000	0.970**	0.758**	0.793**	0.806**
X_5	0.902**	0.220	0.919**	0.970**	1.000	0.801**	0.829**	0.780**
X_6	0.723**	0.251	0.756**	0.758**	0.801**	1.000	0.651**	0.599**
X_7	0.808**	0.065	0.765**	0.793**	0.829**	0.651**	1.000	0.623**
X_8	0.716**	0.288	0.830**	0.806**	0.780**	0.599**	0.623**	1.000

Sumber: SPSS 24

Multikolinearitas dapat diketahui dari nilai koefisien korelasi antar variabel independen(X). Seperti yang tercantum pada Tabel 4.2. Menurut (Gujarati, D. N, 2003), jika nilai koefisien dari variabel independen bernilai kurang dari 0,8 maka variabel tersebut tidak mengalami multikolinearitas. Akan tetapi jika variabel independen bernilai lebih dari atau sama dengan 0,8 maka variabel tersebut mengalami multikolinearitas. Dari hasil Tabel 4.2 pengujian dengan masing-masing nilai P_{value}

disajikan pada lampiran 10 maka disimpulkan bahwa dari 8 variabel independen yang digunakan terdapat 7 variabel yang memiliki nilai lebih dari 0,8 yaitu $X_1, X_3, X_4, X_5, X_6, X_7, X_8$, sedangkan variabel X_2 bernilai kurang dari 0,8. Oleh karena itu, pada penelitian ini hanya Luas Wilayah (X_2) yang tidak bersifat multikolinearitas

4.3 Uji Korelasi Rank Spearman'Rho Antara Variabel Dependen dan Variabel Independen

Uji Korelasi antar variabel penelitian dependen (Y) dan independen (X) juga dilakukan menggunakan nilai Rank Spearman Rho. Sebagai ilustrasi diambil contoh pada variabel Y dengan X_5 , selanjutnya koefisien korelasi antar kedua variabel tersebut diperoleh menggunakan persamaan 22, sehingga diperoleh hasil sebagai berikut:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$r_s = 1 - \frac{6(3365)}{34(34^2 - 1)}$$

$$r_s \frac{636}{1303}$$

$$r_s = 0,486$$

H_0 : Tidak ada korelasi antar variabel Y dan X_5

H_1 : Adanya korelasi antar variabel Y dan X_5

Jika $r_s \geq r_{tabel} = 0,339$, maka H_0 ditolak. dan terima H_1 maka terdapat multikolinearitas.

Dapat dilihat pada hasil persamaan 20, nilai $r_s = 0,486 > r_{tabel} = 0,339$ sehingga H_0 ditolak. maka terdapat multikolinearitas antara Kepadatan Penduduk (Y) dan Jumlah Sekolah Tingkat SMA (X_5).

Tabel 4. Nilai korelasi Spearman'rho antar variabel dependen dan independen

X_1	0.673**
X_2	-0.681**
X_3	0.483**
X_4	0.466**
X_5	0.460**
X_6	0.320

X_7	0.503**
X_8	0.291

Berdasarkan dari hasil pengujian Tabel 4.3, didapatkan nilai korelasi antara variabel dependen (Y) dan variabel independen (X) didapatkan nilai tertinggi yaitu pada variabel $X_1 = 0.673$, $X_2 = -0,681$ dengan nilai P_{value} yang di sajikan pada lampiran 11 sehingga dapat dikatakan nilai variabel independen(X) tersebut memiliki korelasi yang kuat terhadap Y . Dari hasil Tabel 4.3 variabel independen yang memiliki korelasi kuat secara signifikan dengan variabel dependen adalah variabel $X_1, X_2, X_3, X_4, X_5, X_7$, sedangkan variabel Jumlah puskesmas (X_6) dan jumlah kematian (X_8) tidak berkorelasi kuat secara signifikan terhadap variabel Y karena memiliki nilai korelasi yang rendah terhadap Y .

4.4 Uji Equidispersi

Selain uji korelasi, juga diperlukan uji equidispersi untuk mengetahui apakah variabel dependen (Y) memiliki nilai rata-rata dan variansi yang sama atau tidak. Adapun proses pengujian equidispersi dilakukan sebagai berikut:

a) Pengujian Hipotesis

$$H_0 : Var(Y) > E(Y)$$

$$H_1 : Var(Y) = E(Y)$$

b) Jika $Var(Y) > E(Y)$ maka terima H_0 dan tolak H_1 artinya variabel dependen tidak terdapat *equidispersi*

Tabel 5. Hasil Output Uji Equidispersi

Rata - rata	712.62
Variansi	7337423.698

Sumber: SPSS 24

Berdasarkan Tabel 4.4 nilai rata-rata dan variansi variabel dependen (Y) tidak sama sehingga tidak memenuhi asumsi *equidispersi*. Oleh karena asumsi *equidispersi* tidak terpenuhi maka akan dilanjutkan dengan pengujian overdispersi atau underdispersi.

4.5 Pengujian Overdispersi atau Underdispersi

Dalam model regresi poisson terdapat praduga yang harus dipenuhi. Salah satunya adalah anggapan kesamaan antara rata-rata dan perbedaan variabel dependen (Y), yang disebut equidispersion. Bagaimanapun, dalam pemeriksaan sering kali diamati bahwa perubahan lebih menonjol daripada normal atau dapat disebut overdispersi dan jika variabel dependen (Y) menunjukkan variansi lebih kecil dari nilai rata-ratanya, data tersebut dikatakan sebagai underdispersi.

$$\theta_2 = \frac{x^2}{db} > 1; x^2 = \sum_{i=1}^n \frac{(y_i - \mu_1)^2}{\sigma_i}$$

a) Pengujian Hipotesis:

$$H_0 : \chi^2 = 0$$

$$H_1 : \chi^2 \neq 0$$

b) Taraf signifikansi

Selang kepercayaan 95% ($\alpha = 0,05$)

c) Pengambilan keputusan

Jika nilai pearson $\chi^2 > 0$ maka data mengalami kasus Overdispersi maka H_0 ditolak dan H_1 diterima

Tabel 6. hasil Output Memeriksa Overdispersi atau Underdispersi

Godness of Fit			
	Value	df	Value/df
Deviance(D ²)	28553.717	31	921.088
Pearson Chi-Square (χ^2)	1.713×10^{-17}	31	5.527×10^{15}

Sumber: SPSS 24

Dari Tabel 4.5 didapatkan hasil bahwa variabel Kepadatan Penduduk (Y) memiliki nilai Deviance sebesar 921.088 dan nilai *Pearson Chi Square* sebesar 5.527×10^{15} , kedua nilai ini lebih besar dari 1 sehingga dapat disimpulkan terjadi penolakan H_0 sehingga dapat dikatakan data mengalami kasus Overdispersi dan model yang digunakan adalah model Regresi Poisson Tergeneralisasi.

4.6 Estimasi Parameter

Tahapan selanjutnya yang akan dilakukan adalah memodelkan variabel-variabel yang mempengaruhi kepadatan penduduk secara keseluruhan di 34 provinsi di Indonesia menggunakan metode Likelihood. Secara komputasi diperoleh hasil pada Tabel 4.6.

Tabel 7. Estimasi Parameter Model Regresi Poisson Tergeneralisasi

Estimasi Parameter		
Parameter	Estimasi	Sig.
β_0	8.032	< 0.0001
β_1	0.000097	< 0.0001
β_2	-0.000136	< 0.0001

Sumber: SPSS 24

Berdasarkan Tabel 4.6 dapat diperoleh nilai $\beta_0, \beta_1, \beta_2$ sebagai berikut:

$$\beta_0 = 8.032$$

$$\beta_1 = 0.000097$$

$$\beta_2 = -0.000136$$

maka didapatkan model prediksi sebagai berikut:

$$\ln E(y_i|x_i) = \ln(\mu_i) = \eta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$$

$$\mu_i = \exp(x_i^T \beta) = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$$

Persamaan regresi poissonnya adalah:

$$\hat{\mu}_i(X_i) = e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}}$$

$$\hat{\mu}_i(X_i) = e^{(8.032 + (0.000097X_1 - 0.000136X_2))}$$

Dalam model cenderung terlihat bahwa estimasi parameter yang dinilai untuk variabel luas wilayah (X_2) menunjukkan nilai negatif, dan itu dimaknai bahwa hubungan antara variabel ini dan Kepadatan Penduduk (Y) berbanding terbalik. Sedangkan variabel Jumlah Penduduk (X_1) bernilai positif, yang artinya setiap peningkatan satu jiwa penduduk maka tingkat kepadatan penduduk meningkat sebesar $100(1 - e^{0.000097})\% = 9,7\%$, sedangkan luas wilayah (X_2) memiliki hubungan yang berlawanan arah sebesar $100(1 - e^{0.000136})\% = 0,0136\%$, yang artinya semakin sempitnya satu satuan luas wilayah, maka tingkat kepadatan penduduk meningkat sebesar $100(1 - e^{-0.000136})\% = 0,0136\%$.

4.7 Uji Signifikansi Parameter Model

A. Uji Simultan (Uji Statistika G)

Uji Statistika G dilakukan untuk memodelkan faktor-faktor yang diduga memengaruhi kepadatan penduduk (Y). Pemodelan dilakukan dengan terlebih dahulu menguji signifikansi dari nilai estimasi parameter β_1 dan β_2 yang merujuk pada Tabel 4.7 berikut:

a) Pengujian Hipotesis

$$H_0 : \beta_1 = \beta_2 = 0$$

H_1 : paling tidak ada satu β_j yang berpengaruh terhadap model

b) Taraf signifikansi:

Selang kepercayaan 95% ($\alpha = 0.05$)

c) Jika nilai statistik $G < \chi^2_{(\alpha, df)}$ maka terima dan H_0

Jika nilai uji statistik $G \geq \chi^2_{(\alpha, df)}$ maka tolak dan H_0

Tabel 8. Uji Statistika G

Likelihood Ratio	df	Sig.
65266.646	2	< 0.0001

Sumber: SPSS 24

d) Kesimpulan

Dari hasil Tabel 4.7 diketahui bahwa nilai *Likelihood Ratio Chi-Square* sebesar 65266.646. Dapat diketahui nilai statistik uji $G = 65266.646$ dengan nilai $\chi^2(0,05,2) =$

5.9915. Sehingga dapat disimpulkan tolak H_0 dan terima H_1 artinya terdapat salah satu atau lebih variabel independen yang memiliki perbedaan signifikan.

B. Uji Wald

Uji Wald digunakan untuk mengetahui apakah secara parsial variabel independen (X) (Jumlah Penduduk, Luas Wilayah) berpengaruh terhadap variabel dependen (Y) (Kepadatan Penduduk). Statistik uji wald adalah sebagai berikut:

$$W = \left(\frac{\hat{\beta}_j}{SE(\beta_j)} \right)^2$$

Berdasarkan pengukuran uji yang digunakan, nilai W dapat diperoleh sebagai berikut:

$$W_1 = \left(\frac{\hat{\beta}_j}{SE(\beta_j)} \right)^2 = \left(\frac{8.032}{0.0089} \right)^2 = 814455,548$$

a) Pengujian Hipotesis

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

Taraf signifikansi

$$\alpha = 0.05$$

b) Jika $W \geq \chi^2_{(a,df)}$ maka tolak H_0

Tabel 9. Uji Wald

Parameter Estimasi					
Parameter	Estimasi	Std error	Wald	$\chi^2_{(0.05,1)}$	Keputusan
β_0	8.032	0.0089	41.964,108	3,8415	Tolak H_0 dan Terima H_1
β_1	0.000097	0.0000	7.48,073	3,8415	Tolak H_0 dan Terima H_1
β_2	-0.000136	0.0000	1.265,312	3,8415	Tolak H_0 dan terima H_1

Sumber: SPSS 24

Model regresi poisson yang dihasilkan adalah:

$$\hat{\mu}_i(X_i) = e^{\beta_0 + \beta_1 X_i + \beta_2 X_i}$$

$$\hat{\mu}_i(X_i) = e^{(5.441 + (0.000292X_1 - 0.000039X_2))}$$

Parameter $\beta_0, \beta_1, \beta_2$ signifikan pada tingkat signifikansi sebesar 5 % artinya Jumlah Penduduk dan Luas Wilayah memiliki kontribusi terhadap permasalahan Kepadatan Penduduk. Berdasarkan Tabel 4.8 diperoleh tiga koefisien yang signifikan dari hasil dengan uji Wald $> \chi^2_{(a,1)} = 3.8145$. Oleh karena itu, dapat diputuskan variabel Jumlah Penduduk(X_1) dan Luas Wilayah(X_2) mempunyai pengaruh signifikan terhadap Kepadatan Penduduk(Y) dengan nilai AIC sebesar 28786.925.

5. KESIMPULAN

Berdasarkan hasil pemeriksaan dan bahasan dalam penelitian ini, akan diambil kesimpulan sebagai berikut:

- 1) Karakteristik dari kepadatan penduduk di Indonesia lebih didominasi oleh pulau Jawa yaitu provinsi DKI Jakarta yang memiliki nilai Kepadatan Penduduk tertinggi sedangkan untuk yang terendah berada di provinsi Papua Barat.
- 2) Model Regresi Poisson Tergeneralisasi menunjukkan bahwa nilai estimasi parameter model dapat diketahui bahwa tingkat kepadatan penduduk (Y) dan jumlah penduduk (X_1) memiliki hubungan yang bersesuaian secara langsung, maksudnya bahwa untuk setiap penambahan satu penduduk, maka penduduk meningkat sebesar 9,7%, sedangkan luas wilayah (X_2) mempunyai hubungan yang berlawanan, dan itu berarti semakin kecil satu satuan wilayah maka penambahan penduduk bertambah sebesar 0,0136%.
- 3) Dilihat dari model Regresi Poisson Tergeneralisasi cenderung terlihat bahwa variabel yang mempengaruhi kepadatan penduduk di Indonesia tahun 2019 adalah Jumlah Penduduk (X_1), Luas Wilayah (X_2) dengan nilai AIC sebesar 28786.925

DAFTAR PUSTAKA

- [1] Ali, M. I., & Abidin, M. R. (2016). Prosiding Seminar Nasional Lembaga Penelitian Universitas Negeri Makassar "Diseminasi Hasil Penelitian melalui Optimalisasi Sinta dan Hak Kekayaan Intelektual" Pengaruh kepadatan penduduk terhadap intensitas kemacetan lalu lintas di Kecamatan Rappocini Makassar.
- [2] Darnah (2011) 'Handling Overdispersion on Poisson Regression Models with Generalized Poisson', Jurnal Eksponensial, 2(2), pp. 5–10.
- [3] Dewi Sri Susanti, Yuana S. dan N. S. (2018). Buku Analisis Regresi Dan Korelasi.
- [4] Direktorat Jenderal Kependudukan dan Pencatatan Sipil. dukcapil.kemendagri.go.id/berita/baca/809/distribusi-penduduk-indonesia-per-juni-2021-jabar-terbanyak-kaltara-paling-sedikit diakses pada tanggal senin, 21 Februari 16.48
- [5] Fadhillah, Fitriana. 2011. Aplikasi Regresi Binomial Negatif Dan Generalized Poisson Dalam Mengatasi Overdispersion Pada Regresi Poisson (Studi Kasus Data Kemiskinan Provinsi Di Indonesia Tahun 2009). Jakarta.
- [6] Gujarati, D. N. (2003). Basic econometrics. McGraw Hill.
- [7] Harahap, R. I. (2020) 'Pemodelan Metode Regresi Poisson Untuk Anemia Pada Ibu Hamil di Kabupaten Batu Bara'.

- [8] Jihan Sarah Pakpahan. (2018). Faktor-faktor Yang Mempengaruhi Kepadatan Penduduk Sumatera Utara Tahun 2018.
- [9] Nugraha, J. D. (2013) Pengantar Analisis Data Kategorik. Yogyakarta: Deepublish.
- [10] Ruliana, Hendikawati Putjiarti, Agoestanto Arief. 2016. "Pemodelan Generalized Poisson Regression (GPR) untuk campak di kota Semarang tahun 2013."
- [11] Scouse, R. A. (1985) Introduction To Statistical Quality Control. Edition 6t, Plastics and rubber international. Edition 6t. United States of America. doi: 10.2307/2988304.
- [12] Suhardi, Gatot. 2009. "Di Kecamatan Delanggu Kabupaten Klaten Tahun 2006." 0-14.
- [13] Walpole. R. E. 1995. Pengantar Statistika. Jakarta: PT. Gramedia Pustaka Utama.