BAB II

TINJAUAN PUSTAKA

2.1 Regresi Poisson

Regresi Poisson merupakan salah satu jenis analisis regresi yang cocok untuk menentukan hubungan antara variabel terikat (respon) dan variabel bebas (variabel prediktor), dimana variabel respon diasumsikan berdistribusi Poisson. Distribusi Poisson digunakan untuk kejadian-kejadian yang jarang terjadi pada waktu atau wilayah tertentu. Distribusi yang paling sederhana adalah distribusi Poisson, dimana probabilitasnya bergantung pada satu parameter, yaitu mean μ. Fungsi probabilitasnya adalah sebagai berikut:

$$p(y;\mu) = \frac{e^{-\mu}\mu^{y}}{y!},\ untuk\ \ y = 0,1,2,\ldots \tag{2.1}$$

Keterangan :

μ : rata-rata banyaknya sukses dalam selang waktu tertentu atau daerah tertentu.

y : banyaknya sukses dalam selang waktu tertentu atau daerah tertentu

e : nilai konstan 2.7183

Pada metode ini, memiliki asumsi yang harus dipenuhi yaitu asumsi *equidispersi* yaitu $E(Y) = Var(Y)$, dimana $Y \sim Poisson\ (\mu)$. Jika y_i merupakan nilai variabel respon pengamatan ke-i, dimana Y_i bersifat diskrit dan merupakan variabel acak respon, sedangkan x_i adalah nilai variabel prediktor ke-j dari pengamatan ke-i, dimana X_i berjenis diskrit atau kontinue. Secara umum, model regresi Poisson sebagai berikut :

$$y_i = \mu_i + \varepsilon_{i,} \qquad i = 1,2,\ldots,n \tag{2.2}$$

Dimana y_i : nilai pengamatan variabel respon ke-i, dari variabel acak Poisson ; μ_i : rata-rata variabel respon Y yang dipengaruhi variabel prediktor ; $\varepsilon_{i,}$: galat ke-i.

Fungsi penghubung regresi poisson dengan X_i dengan notasi $\mu(X_i, \beta)$,. Dengan demikian fungsi penghubung variabel respon rata-rata μ_i dinyatakan berikut ini :

$$\mu_i = \mu(X_i, \beta) \tag{2.3}$$

$$\mu = \exp(X_i \beta) \tag{2.4}$$

$$= \exp\left[\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_n x_{in}\right] \tag{2.5}$$

Berdasarkan persamaan diatas, maka model umum regresi poisson dinyaakan sebagai berikut :

$$y_i = \exp X_i \beta + \varepsilon_i \tag{2.6}$$

Dimana, X_i yaitu matriks variabel prediktor ukuran $X_i\ n\ x\ (p+1)$, β ialah vektor parameter regresi berdimensi $p+1$, $X_i\beta$ merupakan fungsi linier. $\hat{\beta}$ adalah penduga kemungkinan maksimum β (Nur Pristianing, 2019).

2.1.1 Estimasi Parameter Model Regresi Poisson

Penaksiran Parameter regresi poisson dilakukan menggunakan metode *Maximum Likelihood Estimation* (MLE). Taksiran *maksimum likelihood* untuk parameter β dinyatakan dengan $\hat{\beta}$ merupakan penyelesaian dari turunan pertama dari fungsi likelihood, yaitu :

$$L(\beta) = \prod_{i=1}^{n} P(y_i, \beta) = \frac{\{\prod_{i=1}^{n}[\mu(x_i, \beta)]^{y_i}\} \exp(\sum_{i=1}^{n} \mu(x_i \beta))}{\prod_{i=1}^{n} y_i!} \tag{2.7}$$

Fungsi log natural-likelihoodnya adalah :

$$\ln L(\beta) = \sum_{i=1}^{n} y_i x_i \beta - \sum_{i=1}^{n} \exp(\, x_i \beta) - \sum_{i=1}^{n} \ln y\,! \tag{2.8}$$

2.1.2 Pengujian Parameter Model Regresi Poisson

Pengujian kelayakan regresi Poisson, dilakukan dengan menentukan dua fungsi *likelihood*, dengan melihat hubungan model regresi Poisson. Adapun dua fungsi dimaksud ialah $L(\widehat{\Omega})$ ialah model lengkap untuk nilai *likelihood* dengan memasukkan variabel prediktor, dan $L(\widehat{\omega})$ yaitu model sederhana untuk nilai *likelihood* tanpa memasukkan variabel prediktor.

Uji statistik dan uji parameter pada model regresi Poisson menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dan hipotesis sebagai berikut :

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

H_1 : paling sedikit ada satu $\beta_j \neq 0, . j = 1,2, \dots, k$

Statistik uji (*Likelihood Ratio Test*), model regresi Poisson sebagai berikut ini :

$$G = -2\left[ln\left(\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right)\right] = 2(L(\widehat{\Omega}) - L(\widehat{\omega})) \quad (2.9)$$

Keterangan :

G : Statistik uji parameter secara serentak model regresi Poisson.

$L(\widehat{\Omega})$: fungsi *likelihood* dari model yang melibatkan variabel prediktor

$L(\widehat{\omega})$: fungsi *likelihood* dari model tanpa variabel prediktor

Hasil pengujian, jika tolak H_0 maka $G_{hitung} \geq \chi^2_{v,\alpha}$, dimana v ialah banyaknya parameter di bawah populasi dikurangi dengan banyak parameter di bawah H_0. Parameter yang akan dihasilkan pada model regresi poisson belum tentu berpengaruh signifikan terhadap model. Untuk itu perlu dilakukan pengujian model regresi poisson secara terpisah. Uji parsial menggunakan uji Wald dengan hipotesis berikut ini :

$H_0 : \beta_i = 0$ (tidak ada variabel signifikan)

$H_1 : \beta_i \neq$ (teradapat variabel yang signifikan)

Statistik uji yang digunakan seperti berikut ini :

$$W_i = \left(\frac{\beta_i}{se\ (\beta_i)}\right)^2 \quad (2.10)$$

Kriteria yang digunakan yaitu tolak H_0 jika $W_i > \chi^2_{\alpha,v}$, dimana α yaitu tingkat signifikansi dan v ialah derajat bebas (Badriyah, 2019).

2.2 Overdispersi

Dalam model regresi Poisson ada beberapa asumsi yang harus dipenuhi, yaitu dispersi mean dan variance yang sama. Namun pada data diskrit sering terjadi overdispersi dimana nilai variance lebih besar dari mean. Salah satu penyebab terjadinya overdispersi adalah karena kesalahan spesifikasi model dan kesalahan dalam menentukan fungsi link. Jadi, ini tidak baik untuk model regresi Poisson. Oleh karena itu, fenomena overdispersi dapat dituliskan sebagai:

$$Var(Y) > E(Y) \tag{2.11}$$

Terdapatnya *overdispersi* dideteksi melalui rasio antara devians dan derajat bebasnya. Apabila rasio yang dihitung menghasilkan nilai yang lebih besar dari satu, maka ada dua cara untuk mendeteksi *overdispersi* diantaranya :

1. *Deviance*

$$\theta_1 = \frac{D}{db} > 1; D = 2\sum_{i=1}^{n}\left\{y_i \log\frac{y_i}{\mu_i}\right\} \tag{2.12}$$

2. *Pearson Chi-Square*

$$\theta_2 = \frac{\chi^2}{db} > 1; \chi^2 = \sum_{i=1}^{n}\frac{(y_i-\mu_i)^2}{\sigma_i} \tag{2.13}$$

Keterangan :

y_i : nilai peubah dari pengamatan ke-i

μ_i : parameter bagi respon rata-rata ke-i

σ_i : parameter bagi respon ke-i

db : $n - k$

k : banyaknya parameter termasuk konstanta

n : banyaknya pengamatan

θ : Parameter

Jika, θ_1 dan θ_2 memiliki nilai lebih dari 1, maka terjadi *overdispersi* data (Greis, 2017).

2.3 Multikolinieritas

Multikolinieritas terjadi karena adanya hubungan diantara variabel prediktor dalam model regresi. Adanya multikolinieritas diketahui dengan menghitung nilai *Varians Inflation Factor* (VIF). Apabila nilai VIF pada variabel prediktor lebih besar dari 10, maka hal tersebut menunjukkan adanya multikolinieritas. Untuk menghitung nilai VIF sebagai berikut :

$$VIF = \frac{1}{1-R_j^2}, \;\; j = 1,2,\dots,p, \tag{2,14}$$

Keterangan

R_i^2 : Koefisien penentuan antara variabel prediktor (x) (Amalia, 2022).

2.4 Regresi *Zero-Inflated Poisson* (ZIP)

Salah satu penyebab *overdispersi* yaitu varians lebih besar dari mean sehingga model regresi poisson tidak efektif digunakan. Model regresi ZIP merupakan salah satu metode digunakan untuk mengatasi masalah *overdispersi* disebabkan pada data diskrit. Dalam model ZIP terdapat 2 paraneter yang ditaksir perlu diestimasi. Keadaan pertama yaitu *zero state* terjadi dengan probabilitas ω_i dan menghasilkan hanya observasi bernilai nol. Sementara keadaan kedua yaitu *Poisson state* terjadi dengan probabilitas $1-\omega_i$ dan berdistribusi Poisson dengan mean μ_i.

Untuk setiap pengamatan Y_i bebas satu sama lain i n = 1, 2,..., dan

$$Y_i = \begin{cases} 0 & dengan\, probbilitas\;\; \omega_i \\ Poisson\;(\mu_i) & dengan\, probabilitas\;(1-\omega_i) \end{cases}$$

Kemudian;

$$\mathrm{P}(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu} & jika\; y_i = 0 \\ (1-\omega_i)\frac{e^{-\mu}\mu^{y_i}}{y!} & jika\; y_i = 1,2,\dots;0 \le \omega_i \le 1 \end{cases} \tag{2.15}$$

Lambert menyarankan model gabungan untuk μ dan ω adalah :

$$\ln(\mu_i) = x_i\beta \;\text{ dan }\; \log(\omega) = \ln\left(\frac{\omega_i}{1-\omega_i}\right) = x_i\gamma \tag{2.16}$$

Dimana X_i : Matriks variabel prediktor, $\beta\ dan\ \gamma$: vektor parameter yang akan dinilai. ω : probabilitas pengamatan bernilai nol. Dengan demikian, untuk rumus model regresi ZIP adalah sebagai berikut:

$$\ln \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} \,; i = 1,2, \dots, n \tag{2.17}$$

$$\text{logit}\, \omega_i = \gamma_0 + \gamma_1 x_{i1} + \cdots + \gamma_k x_{ik} \,; i = 1,2, \dots, n \tag{2.18}$$

Dimana, $\beta_{(k+1)x1}$ dan $\gamma_{(k+1)x1}$ merupakan vektor parameter model regresi ZIP. Pendugaan parameter ini, diperoleh dengan metode *Maximum Likelihood Estimation* (Adeliana,2017).

2.4.1 Penaksiran Parameter Model Regresi ZIP.

Metode yang digunakan untuk mengestimasi parameter ZIP adalah *Maximum Likelihood Estimation* biasa digunakan untuk menaksir parameter yang diketahui fungsi densitasnya. Untuk menaksir parameter regresi ZIP digunakan metode *Maksimum Likelihood Estimation* (Samson, 2022) :

$$\mu_i = \exp(\, x_i\, \beta) \tag{2.19}$$

$$\omega_i = \frac{e^{x_i \gamma}}{1+e^{x_i \gamma}} \tag{2.20}$$

$$(1 - \omega_i) = \frac{1}{1+e^{x_i \gamma}} \tag{2.21}$$

Selanjutnya, mensubstitusikan persamaan diatas :

$$P(Y = y_i) = \begin{cases} \dfrac{e^{x_i \gamma} + (\exp(-e^{x_i \beta}))}{1+\exp(-e^{x_i \beta})} & ,\ untuk\ \ y_i = 0 \\ \dfrac{\frac{1}{1+e^{x_i \gamma}}\left(exp\left((-e^{x_i \beta}) + (x_i \beta)^{y_i}\right)\right)}{y_i} & ,untuk\ y_i > 0 \end{cases} \tag{2.22}$$

Untuk fungsi *likelihood* sebagai berikut :

$$L(\gamma, \beta | y_i) = \begin{cases} \prod_{i=1}^{n} \dfrac{e^{x_i \gamma} + (\exp(-e^{x_i \beta}))}{1+e^{x_i^T \gamma}} & ,untuk\ y_i = 0 \\ \dfrac{\prod_{i=1}^{n} \frac{1}{1+e^{x_i \gamma}}\left(exp\left((-e^{x_i \beta}) + (x_i \beta)^{y_i}\right)\right)}{\prod_{i=1}^{n} y_i!} & ,untuk\ \ y_i > 0 \end{cases} \tag{2.23}$$

2.4.2 Pengujian Parameter Model Regresi ZIP

Pengujian parameter model ZIP menggunakan metode penaksir *Maximum Likelihood Ratio Test* (MLRT). Untuk masing-masing hipotesis dan statistik *likelihood*, sebagai berikut :

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = \gamma_1 = \gamma_2 = \cdots \gamma_1 = 0$$

$$H_1 : paling\ sedikit\ ada\ satu\ \beta_r \neq 0{,}0 < r < k + 1$$

Keterangan :

$k + 1$: Jumlah parameter

β_r : Parameter model log ke-r

γ_r : Parameter model logit ke-r

Menurut *Lestari* (2009), untuk melakukan perhitungan uji statistik pengujian model ZIP. Maka, untuk kesesuaian uji model berikut ini :

$$G = -2\left[\frac{L(y;\hat{\omega})}{L(y;\hat{\Omega})}\right] \tag{2.24}$$

Selanjutnya, untuk fungsi *likelihood* :

$$L(\Omega)\prod_{i=1}^{n} f(y_i;\beta,\gamma) = \prod_{i=1}^{n}\left[\left(\frac{1}{1+\exp(x_i\gamma)}\right)(\exp(x_i\gamma))^{z_i}\left(\frac{\exp\left(-e^{x_i\beta}\right)(\exp(x_i\beta)^{y_i}}{y_i!}\right)^{1-z_i}\right] \tag{2.25}$$

Kemudian, untuk fungsi *likelihood* jika H_0 benar:

$$L(\omega)\prod_{i=1}^{n} f(y_i;\beta_0,\gamma_0) = \prod_{i=1}^{n}\left[\left(\frac{1}{1+\exp(\gamma_0)}\right)(\exp(\gamma_0)^{z_i}\left(\frac{\exp(-e^{\beta_0})(\exp(\beta_0)^{y_i}}{y_i!}\right)^{1-z_i}\right] \tag{2.26}$$

Dimana, (Ω): $\Omega\{\beta,\gamma\}$ merupakan himpunan parameter dibawah populasi (ω): ω: $\{\beta_0,\gamma_0\}$ merupakan himpunan parameter jika H_0 benar $\beta\ dan\ \gamma$:parameter yang ditaksir.

Statistik uji G mengikuti nilai dari distribusi *chi-square* yaitu dengan kriteria tolak H_0 apabila $G > \chi^2_{\alpha,2v}$, artinya paling sedikit ada satu parameter yang memberikan pengaruh signifikan terhadap variabel respon.

Setelah melakukan pengujian serentak parameter, seterusnya yaitu dilakukan pengujian parameter secara parsial yang bertujuan untuk melihat pengaruh signifikan dimasing-masing parameter β dan γ. Uji parameter parsial dilakukan menggunakan uji *Wald* seperti berikut ini :

a. Pengujian untuk parameter β

Untuk hipotesis yang digunakan adalah sebagai berikut

$H_0 : \beta_i = 0$ (tidak ada variabel signifikan)

$H_1 : \beta_i \neq 0$ (teradapat variabel yang signifikan)

Statistik uji yang digunakan seperti berikut ini :

$$W_i = \left(\frac{\beta_i}{se\ (\beta_i)}\right)^2$$

Statistik uji *wald* digunakan mengikuti distribusi *chi-square*, sehingga kriteria keputusannya adalah tolak H_0 jika $W_i > \chi^2_{\alpha,1}$ artiny parameter yang berpengaruh signifikan terhadap variabel respon.

b. Pengujian untuk parameter γ

Untuk hipotesis yang digunakan adalah sebagai berikut

$H_0 : \beta_i = 0$ (tidak ada variabel signifikan)

$H_1 : \beta_i \neq 0$ (teradapat variabel yang signifikan)

Statistik uji yang digunakan seperti berikut ini :

$$W_i = \left(\frac{\gamma_i}{se\ (\gamma_i)}\right)^2$$

Statistik uji wald digunakan mengikuti distribusi *chi-square*, sehingga kriteria keputusannya adalah tolak H_0 jika $W_i > \chi^2_{\alpha,1}$ artinya terdapat parameter yang berpengaruh signifikan pada variabel respom (Lestari, 2009).

2.5 *Geographically Weighted Zero-Inflated Poisson Regression* (GWZIPR)

Metode Metode *Geographically Weighted Zero-Inflated Poisson Regression* merupakan model pengembangan dari *Geographically Weighted Regression* (GWR) dan *Zero-Inflated Poisson* (ZIP) yang pertama kali diusulkan oleh Kalogirou pada tahun 2015. GWZIPR adalah metode alternatif untuk menyelesaikan data diskrit yang melibatkan banyak angka nol. efek spasial. Tujuan penggunaan GWZIPR adalah untuk mempelajari faktor-faktor yang mempengaruhi data kasus Covid-19 dan variabel prediktor yang mempengaruhi setiap kabupaten/kota di Sumatera Utara karena melibatkan efek spasial.

Pemodelan data spasial selalu melibatkan matriks bobot spasial. Dampak spasial pada data dapat disebut sebagai kesalahan korelasi antara lokasi ketergantungan spasial dan keragaman spasial (heterogenitas). Penentuan matriks bobot dan efek spasial dijelaskan lebih lanjut sebagai berikut:

2.5.1 Penentuan *Bandwidth* dan Pembobot Optimum

Suatu bobot memiliki peran penting dalam data spasial, untuk mewakili pengambilan lokasi setiap data. Sebuah lokasi dapat diwakili oleh titik koordinat, dimana lokasi terdekat harus menunjukkan hubungan sama dan jika lokasi berjauhan menunjukan heterogenitas spasial, ditunjukkan oleh matriks berbobot $w(u_i, v_j)$ merupakan jarak *Euclidean* dengan rumus berikut ini :

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{2.27}$$

Pembobotan untuk model GWR disetiap lokasi dapat ditentukan menggunakan fungsi kernel Gaussian, sebagai berikut:

$$w_{ij}\left(u_i, v_j\right) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right) \tag{2.28}$$

Dimana, d_{ij} : jarak *Euclidean* antar lokasi ke-i dan lokasi ke-j dan h : *bandwidth* atau parameter pemulus optimum (Amaliana, 2018).

Dalam melakukan pemilihan *bandwidth* optimum akan jadi sangat penting karena *bandwidth* akan mempengaruhi model data yang akan mengatur varians dan bias pada model. Salah satu untuk menentukan *bandwidth* optimum pada model GWR yaitu dengan *cross-validation* (CV). Nilai CV terhadap parameter yang diperlukan dari bobot, adapun fungsi yang dipilih akan memberikan nilai parameter sesuai dan menghasilkan nilai CV minimum. Salah satu aspek untuk menentukan *bandwidth* optimum dengan *Cross Validation* (CV) atau validitas silang berikut ini :

$$CV = \sum_{i-1}^{n} [y_i - \hat{y}_{\neq i}(h)]^2 \tag{2.29}$$

Dimana, $\hat{y}_{\neq i}(h)$ adalah penduga dari y_i, jika lokasi (u_i, v_i) adalah tidak termasuk dalam estimasi dan n adalah banyaknya lokasi. Nilai CV minimum dapat diperoleh ketika bandwitdh optimal digunakan (Maulidiyah, 2018).

2.5.2 Dependensi Spasial

Dependensi spasial dilakukan untuk menunjukkan bahwa keberadaan lokasi pengamatan antara satu lokasi dipengaruhi oleh lokasi lainnya. Salah satu uji, untuk mengetahui apakah hubungan antara lokasi satu dengan lokasi lainnya setiap variabel dapat diketahui dengan uji *moran's I*. *Moran's I* adalah metode untuk menentukan apakah ada wilayah ketergantungan spasial. Uji statistik *Moran's I* selalu digunakan untuk autokorelasi spasial, dimana penggunaannya diukur dengan korelasi antara pengamatan di satu lokasi dengan lokasi lainnya. Uji *moran's I* didefinisikan dengan rumus berikut :

$$I = \frac{n \sum_{i=1}^{n} \sum_{p=1}^{n} w_{ip}(x_i - \bar{x})(x_p - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{2.30}$$

Dimana, I adalah Indeks Moran, n adalah banyaknya pengamatan, x_i adalah nilai variabel prediktor pada lokasi i, x_p adalah nilai variabel prediktor pada lokasi p, $p = i + 1$, $\bar{x}$ adalah rata-rata variabel prediktor dan w_{ip} adalah elemen matriks pembobot antar lokasi i dan p.

Nilai Indeks Moran adalah -1 hingga 1. Jika nilai *Moran's* $(I > 0)$, artinya ada autokorelasi spasial pada lokasi pengamatan yang berdekatan. Sedangkan, apabila nilai Moran $(I < 0)$, artinya adanya autokorelasi spasial pada lokasi yang berdekatan, namun cenderung berbeda. Kemudian, jika nilai Moran $(I = 0)$, artinya bahwa tidak ada autokorelasi spasial pada data.

Untuk melakukan uji dependensi spasial hipotesis digunakan sebagai berikut :

$H_0 : I = 0$ (adanya dependensi spasial)

$H_1 : I \neq 0$ (tidak ada dependensi spasial)

Uji statistik dependensi spasial menggunakan Moran's I :

$$Z_{hit} = \frac{I - E(I)}{\sqrt{Var(I)}} \quad (2.31)$$

Keterangan :

Z_{hit} : statistik uji Moran's I

I : Indeks Moran's I

$E(I)$: ekspektasi Moran's I besarnya $-\frac{1}{(n-1)}$

Var (I) : variansi dari Moran's I

Kriteria keputusan, jika nilai $|Z_{hit}| > Z_{\alpha/2}$ atau nilai $p - value > \alpha$ maka H_0 gagal ditolak, artinya tidak terdapat dependensi spasial (Alfi, 2021).

2.5.3 Heterogenitas Spasial

Heterogenitas spasial ini terjadi karena perbedaan karakteristik observasi antara satu lokasi dan lokasi yang lain. Adanya perbedaan titik lokasi pengamatan dengan lokasi lain menyebabkan keragaman spasial. Heterogenitas spasial dilakukan untuk melihat karakteristik suatu lokasi observasi. Pengaruh yang terjadi pada lokasi pengamatan akibat adanya heterogenitas spasial ialah membuat parameter model regresi yang berbeda-beda secara spasial. Pada kali ini, uji heterogenitas spasial yang digunakan adalah menggunakan uji *Breusch Pagan* dengan hipotesis berikut ini :

$H_0 : \sigma^2(u_i, v_i) = \cdots = \sigma^2 u_m, v_n = \sigma^2$ (variansi antar lokasi sama)

$H_1 : \sigma^2(u_i, v_i) \neq \sigma^2$ (variansi antar lokasi berbeda)

Maka, statistik dari uji *Breusch Pagan* berikut ini :

$$BP = \left(\frac{1}{2}\right) f^T Z (Z^T Z)^{-1} Z^T f \sim x^2_{(k)} \tag{2.32}$$

dengan vektor f berukuran n x 1 adalah

$$f = \left[\left(\frac{e_i^2}{\sigma^2}\right) - 1, \dots, \left(\frac{e_n^2}{\sigma^2}\right) - 1\right]^T \tag{2.33}$$

keterangan :

i : $1,2,\dots,n; n =$ banyak lokasi pengamatan

k : banyak variabel independen

Z : matriks berdimensi $n\ x(k+1)$ yang berisi variabel independen yang telah distandarisasikan untuk setiap pengamatan .

σ^2 : ragam residual (e_i) dari modelregresi poisson

e_i : residual untuk pengamatan ke-i dari model regresi poisson

Adapun kriteria yang dipakai yaitu tolak H_0 jika nilai $BP > \chi^2_{(\alpha,k)}$ artinya memiliki variansi atau ragam pada lokasi berbeda (Adeliana, 2017).

Model GWZIPR ialah salah satu bentuk lokal dari model regresi *Zero-Inflated Poisson* (ZIP) akan menghasilkan penaksir parameter model bersifat lokal setiap lokasi. Untuk setiap pengamatan pada variabel respon diambil dari lokasi berbeda (u_i, v_j) dan saling bebas :

$$y_i \sim \begin{cases} 0, & dengan\ peluang\ \omega_i \\ Poisson\ (\mu), & dengan\ peluang\ (1-\omega_i) \end{cases}$$

Dalam model GWZIPR variabel dependen Y_i memiliki perbedaan peluang untuk $y_i = 0$ dan $y_i > 0$, berikut :

$$P(Y = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu_i} & ,untuk\ y_i = 0 \\ \frac{(1-\omega_i)e^{-\mu_i}\mu_i{}^{y_i}}{y_i!} & ,untuk\ y_i > 0 \end{cases} \tag{2.34}$$

Yang mana,

$$\mu_i = e^{x_i\beta(u_i,v_j)} \quad dan \quad \omega = \frac{e^{x_i\gamma(u_i,v_j)}}{1-e^{x_i\gamma(u_i,v_j)}} \tag{2.35}$$

Dalam hal ini, $\beta(u_i, v_j)$ dan $\gamma(u_i, v_j)$ adalah parameter yang terletak dalam (u_i, v_j), X adalah matriks kovariat dikaitkan dengan probabilitas *zero state* $(y_i = 0)$ dan mean pada *Poisson state* $(y_i > 0)$.

2.6.1 Estimasi Parameter model GWZIPR

Estimasi parameter dalam GWZIPR adalah *Maximum Likelihood* (ML) dengan fungsi *likelihood* berikut :

$$\ln L(\gamma(u_i, v_j), \beta(u_i, v_j)) = \begin{cases} \sum_{i=1}^{n}\left(\ln\left(e^{x_i\gamma(u_i,v_j)} + e^{-e^{x_i\beta(u_i,v_j)}}\right) - \ln\left(1+e^{x_i\gamma(u_i,v_j)}\right)\right) w_{ij}(u_i, v_j)\ , y_i = 0 \\ \sum_{i=1}^{n}\left(\ln\left(e^{x_i\beta(u_i,v_j)} + y_i x_i\beta(u_i, v_j)\right) - \ln\left(1+e^{x_i\gamma(u_i,v_j)}\right) - \ln(y_i!)\right) w_{ij}(u_i, v_j), y_i > 0 \end{cases} \tag{2.36}$$

Berdasarkan penduga parameter model GWZIPR bisa dilakukan menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan fungsi *likelihood* dan *log-natural likelihood* dengan persamaan sebagai berikut ini:

$$L(\beta(u_i, v_j), \gamma(u_i, v_j)) = \prod_{i=1}^{n} \frac{\exp(x_i\gamma) + (\exp(-\exp(x_i\beta)))}{1+\exp(x_i\gamma)} + \frac{\prod_{i=1}^{n} \frac{1}{\exp(x_i\gamma)} (\exp(x_i\beta)))(-\exp(x_i\beta))^{y_i}}{\prod_{i=1}^{n} y_i!} \tag{2.37}$$

$$
\begin{aligned}
\ln\left(\beta(u_i, v_j), \gamma(u_i, v_j)\right) \\
&= \sum_{\mathrm{i=1}}^{\mathrm{n}} \ln(\exp(x_i\gamma) + (\exp(-\exp(x_i\beta)))w_{ij}\left(u_i, v_j\right) \\
&- \sum_{i=1}^{n} \ln(\exp(x_i\gamma))w_{ij}\left(u_i, v_j\right) \\
&+ \sum_{i=1}^{n} (y_i x_i \beta - \exp(x_i\beta)\, w_{ij}\left(u_i, v_j\right) \\
&- \sum_{i=1}^{n} \ln(y_i!)\, w_{ij}\left(u_i, v_j\right) \qquad (2.38)
\end{aligned}
$$

keterangan : $w_{ij}\left(u_i, v_j\right)$: bobot untuk wilayah ke-i, $i = 1,2,$ ...,n

Penggunaan metode GWZIPR dapat dianalisis untuk seluruh wilayah atau dapat juga di analisis per wilayah (Ismah, 2020).

2.6.2 Pengujian Parameter Model GWZIPR

Pengujian serentak model GWZIPR dilakukan untuk mengetahui tingkta signifikansi terhadap parameter $\beta(u_i, v_j)$ dan $\gamma(u_i, v_j)$. Maka hipotesis yang digunakan untuk pendugaan parameter model GWZIPR berikut ini.

H_0: $\beta_j(u_i, v_i) = \cdots = \beta_k(u_i, v_i) = \gamma_j(u_i, v_i) = \cdots = \gamma_k(u_i, v_i) = 0$

H_1: paling tidak terdapat satu $\beta_j(u_i, v_i) \neq 0$ atau $\gamma_j(u_i, v_i) \neq 0$

Statistik uji yang digunakan ialah dengan uji rasio *likelihood* (devians), yang mana nilai devians terbentuk dari parameter model lengkap $(\Omega = (\beta(u_i, v_i), \gamma(u_i, v_i))$ dan model intersep $(\omega = (\beta(u_i, v_i), \gamma(u_i, v_i))$). Nilai devians terbentuk dari dua ln $likelihood$berikut ini :

$$G = -2\left[\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right] \sim \chi^2_{n-k+1}$$

G berdistribusi dengan *chi-square* dengan derajat bebas $(n - (k + 1))$, dimana n yaitu banyaknya populasi dan k yaitu banyaknya variabel prediktor. Tolak H_0 apabila nilai $G > \chi^2_{(n-(k+1))\alpha}$.

Langkah selanjutnya yaitu pengujian parameter secara parsial, dilakukan untuk mengetahui faktor-faktor yang berpengaruh signifikan terhadap jumlah kasus dengan model GWZIPR, dengan masing-masing parameter $\beta(u_i, v_i)$ dan masing-masing parameter $\gamma(u_i, v_i)$, di setiap lokasi pengamatan.

Hipotesis yang digunakan adalah berikut ini :

$H_0 : \beta_j(u_i, v_i) = 0$; $H_1 : \beta_j(u_i, v_i) \neq 0$

$H_0 : \gamma_j(u_i, v_j) = 0$; $H_1 : \gamma_j(u_i, v_j) \neq 0$

dengan $j = 1,2, \dots 5$; $i = 1,2, \dots ,34$

Pengujian parsial parameter $\beta(u_i, v_j)$ menggunakan uji statistik uji Wald (W) adalah :

$$W = \frac{\beta(u_i,v_j)}{se(\beta(u_i,v_j))} \sim Z_{(0,1)} \tag{2.39}$$

Dimana, kriteria keputusan, tolak H_0 Jika $W > Z_{\alpha/2}$, artinya bahwa parameter ke-j pada lokasi ke-i pada garis lintang dan bujur (u_i, v_j) berpengaruh signifikan terhadap model GWZIPR.

Pengujian parsial parameter $\gamma(u_i, v_j)$ menggunakan uji statistik Uji Wald (W) adalah :

$$W = \frac{\gamma(u_i,v_j)}{se(\gamma(u_i,v_j))} \sim Z_{(0,1)} \tag{2.40}$$

Dimana, kriteria keputusan tolak H_0 Jika $W > Z_{\alpha/2}$, artinya bahwa parameter ke-j pada lokasi ke-i pada garis lintang dan bujur (u_i, v_j) berpengaruh signifikan terhadap model GWZIPR (Faricha, 2016).

2.6 Penyakit Covid-19.

Covid-19 merupakan penyakit menular yang disebabkan oleh virus corona baru yang termasuk dalam keluarga Severe Acute Respiratory Syndrome (SARS) dan Middle East Respiratory Syndrome (MERS). Meski berasal dari keluarga yang sama, namun Covid-19 menular dengan sangat cepat dan berujung pada kematian. Covid-19 merupakan virus RNA dengan ukuran partikel 120-160 nm yang menyerang saluran pernapasan manusia.

Sebelum wabah terjadi, ada enam jenis Covid-19 yang terinfeksi secara global, yaitu alphacoronavirus 229E, alphacoronavirus NL63, betacoronavirus OC43, betacoronavirus HKU1, virus pernapasan akut parah (SARS-CoV), dan Middle East Respiratory Syndrome Virus Corona (MERS). Virus corona atau Covid-19 termasuk dalam genus betacoronavirus yang mirip dengan SARS. Salah satu virus penyebab infeksi parah, Covid-19, ditentukan oleh respon imun tubuh.Sistem imun yang tidak kuat dalam merespon tingkat keparahan infeksi, dan disisi lain respon imun berlebihan juga ikut andil dalam kerusakan jaringan.

Masa inkubasi Covid-19 adalah 5-6 hari, berkisar antara 1 hingga 14 hari. Covid-19 menyebar melalui droplet dari orang yang bergejala ke orang lain di sekitarnya. Penularan juga bisa terjadi melalui benda-benda di sekitar yang terkontaminasi oleh droplet orang yang terinfeksi. Gejalanya ringan, seperti demam, kelelahan, dan batuk kering.

Ketika jumlah kasus Covid-19 meningkat, risiko kesehatan masyarakat global kemungkinan besar akan mendorong karakterisasi variances of interest (VOI) dan variances of concern (VOC). Varians of interest (VOC) meliputi Alpha (B.1.1.7), Beta (B1.351), Gamma (p.1), Delta (B.1617.2), dan Omicron (B.1.1.529). Mengurangi terjadinya mutasi yang berdampak negatif terhadap kesehatan masyarakat memerlukan strategi dan tindakan yang direkomendasikan oleh Organisasi Kesehatan Dunia (WHO) untuk menegakkan protokol kesehatan yang ketat selama bepergian (Adityo, 2020).

2.7 Faktor Pengaruh Jumlah Kasus Covid-19.

Beberapa faktor yang mempengauhi jumlah kasus Covid-19, yang memicu luasnya sebaraan penyakit covid-19 sebagai berikut :

a) Kepadatan Penduduk

Kepadatan penduduk adalah ukuran sebaran yang menunjukkan jumlah penduduk per kilometer persegi suatu wilayah. Sumatera Utara merupakan provinsi terpadat keempat di Indonesia, setelah Jawa Barat, Jawa Timur, dan Jawa

Tengah. Kepadatan penduduk menjadi salah satu faktor yang berkontribusi terhadap peningkatan penyebaran Covid-19 di suatu daerah. Kepadatan penduduk mempunyai peranan penting dalam penyebaran Covid-19, karena kepadatan penduduk yang tinggi di perkotaan dapat menyebabkan penularan penyakit yang lebih tinggi (BPS Sumut, 2022).

b) Kemiskinan

Kemiskinan dipandang sebagai ketidakmampuan ekonomi untuk memenuhi kebutuhan pokok makanan dan non-makanan yang diukur dengan pengeluaran. Tingkat kemiskinan juga mempunyai dampak yang signifikan terhadap jumlah kasus Covid-19, karena individu dan kelompok miskin yang tidak memiliki layanan kesehatan dapat menerima informasi yang salah sehingga mengabaikan peringatan kesehatan masyarakat. Pandemi covid-19 menyebabkan orang jatuh pada kemiskinan, karena aktivitas ekonomi diluar ruangan dibatasi untuk menghindari bertambahnya jumlah kasus Covid-19 (BPS Sumut, 2022).

c) Penduduk Lansia

Menurut Dinas Kesehatan Provinsi Sumut, sebaran jumlah kasus Covid-19 dipengaruhi oleh proporsi penduduk lanjut usia, akibat menurunnya imunitas seiring bertambahnya usia. Salah satu kelompok yang paling berisiko tertular Covid-19 adalah lansia, dimana mereka yang berusia di atas 65 tahun memiliki risiko 2,6 kali lebih tinggi tertular Covid-19 dibandingkan lansia. Faktor usia yang menjadikan masyarakat paling rentan terkena Covid-19 adalah karena lansia memiliki penyakit penyerta yang meningkatkan risiko tertular Covid-19. (Dinas Kesehatan Sumut, 2021).

d) Rumah Tangga Dengan Sanitasi Layak

Rumah tangga adalah seseorang atau sekelompok orang yang mendiami sebagian atau seluruh bangunan fisik/sensus, biasanya tinggal bersama dan mengelola makanan dari satu dapur. Kebersihan yang buruk memberikan dampak yang signifikan terhadap seluruh aspek kehidupan, salah satunya adalah penyebaran penyakit yang tentu saja semakin meningkat (BPS Sumut, 2022).

2.8 Penelitian Terdahulu

1. Penelitian tentang penyebaran jumlah kasus Covid-19, telah dilakukan oleh Ilham Faishal (2020) dengan Judul jurnal, "**Pemodelan Jumlah Kasus Covid-19 Menggunakan *Geeoraphically Weighted Regression***". Hasil penelitian yaitu bahwa faktor yang berpengaruh pada penyebaran Covid-19 disebagian wilayah Propinsi Jawa Barat dipengaruhi oleh tingkat kemiskinan.
2. Penelitian sebelumnya mengenai kasus Covid-19, dilakukan oleh Heinrich, Mangapul dan Rudy (2021) dengan judul jurnal **"Analisis Sebaran Spasial Tingkat Kejadian Kasus Covid-19 dengan Metode Kernel Density di Kota Ambon**". Hasil penelitian yaitu bahwa sebaran spasial tingkat kejadiaan positif Covid-19 terkonsentrasi pada wilayah-wilayah tertentu mengikuti jumlah kepadatan penduduk dan faktor lingkungan lain. Dan hubungan kepadatan penduduk berkorelasi kuat.
3. Penelitian sebelumnya telah dilakukan oleh Rahmat (2020) dengan judul jurnal "**Analisis Jumlah Kasus Malaria di Wilayah Sumatera Menggunakan *Geographically Weighted Zero-Inflated Poisson Regression***". Berdasarkan hasil penelitiannya, bahwa setiap daerah memiliki model khas berdasarkan variabel signnifikan dan pemetaan tidak membentuk pola penyebaran dalam kelompok tersebut karena letak geografis yang berdekatan dan menghasilkan pemetaan sebaran faktor yang mempengaruhi jumlah kasus malaria dan membagi tiga kelompok variabel signifikan dengan penggunaan GWZIPR merupakan metode terbaik daripada metode ZIP regression.
4. Penelitian mengenai metode GWZIPR telah dilakukan oleh Amaliana (2018) dengan judul jurnal "***Comparison of Two Functions in Geographically Weighted Zero-Inflated Poisson Regression on Filariasis Data***". Berdasarkan hasil temuannya, bahwa nilai devians model, diketahui bahwa model GWZIPR menggunakan kernel Gaussian tetap lebih baik daripada GWZIPR menggunakan kernel *Bi-square* tetap.
5. Penelitian sebelumnya mengenai metode GWZIPR telah dilakukan oleh Ismah,dkk (2020) dengan judul jurnal ***"Geographically Weighted Zero-Inflated Poisson Regression* (GWZIPR) Dengan Pembobot *Fixed Bisquare Kernel* Pada Kasus Difteri Di Indonesia**". Hasil penelitiannya,

bahwa pembobot dengan *fungsi kernel bisquare* menghasilkan signifikan antar variabel bebas. Hal tersebut disebabkan terdapat hubungan yang kua dan apabila model tidak dilibatkan maka variabel persentase jumlah kepadatan penduduk, persentase jumlah orang yang divaksinasi mempengaruhi jumlah imunisasi dapat mengurangi jumlah kematian yang diakibatkan penyakit difteri.

2.9 Konsep Wahdatul Ulum

2.9.1 Pandangan Al-Qur’an dan Hadist tentang Wabah Covid-19

Sebagaimana, hadits shahih berkaitan dengan wabah (Covid-19) yang sedang melanda dunia yaitu hadits yang diriwayatkan oleh Imam Bukhori dan Muslim yaitu:

قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ الطَّاعُونُ آيَةُ الرِّجْزِ ابْتَلَى اللَّهُ عَزَّ وَجَلَّ بِهِ
نَاسًا مِنْ عِبَادِهِ فَإِذَا سَمِعْتُمْ بِهِ فَلَا تَدْخُلُوا عَلَيْهِ وَإِذَا وَقَعَ بِأَرْضٍ وَأَنْتُمْ بِهَا فَلَا تَفِرُّوا مِنْهُ

Artinya : “ Rasulullah SAW bersabda : Tha’un (wabah penyakit menular) adalah suatu peringatan dari Allah SWT untuk menguji hamba-hamba-Nya dari kalangan manusia. Maka apabila kamu mendengar penyakit itu berjangkit di suatu negeri, janganlah kamu masuk ke negeri itu, dan apabila wabah itu berjangkit di negeri tempat kamu berada, maka janganlah pula kamu lari daripadanya.”(HR.Bukhari dan Muslim).

Sebagaimana dijelaskan dalam hadis di atas, mewabahnya penyakit Tarn yang memiliki ciri-ciri yang sama dengan Covid-19 yang melanda berbagai belahan dunia ini merupakan sisa dari azab yang terjadi pada masa Nabi Muhammad SAW. nabi. Ibnu Hajar menjelaskan dalam kitab “Fath Al-Barri” bahwa azab yang dimaksud adalah bagi orang-orang kafir dan orang-orang yang maksiat, sedangkan jika seorang mukmin meninggal karena wabah Covid-19, termasuk syahid, maka wabah ini adalah rahmat bagi mukmin.

Sebagaimana, firman Allah Swt tentang kajian Covid-19 yang terkandung dalam QS. Al-Baqarah : 26.

۞ اِنَّ اللّٰهَ لَا يَسْتَحْيٖٓ اَنْ يَّضْرِبَ مَثَلًا مَّا بَعُوْضَةً فَمَا فَوْقَهَاۗ فَاَمَّا الَّذِيْنَ اٰمَنُوْا فَيَعْلَمُوْنَ اَنَّهُ الْحَقُّ مِنْ رَّبِّهِمْۚ وَاَمَّا الَّذِيْنَ كَفَرُوْا فَيَقُوْلُوْنَ مَاذَآ اَرَادَ اللّٰهُ بِهٰذَا مَثَلًاۘ يُضِلُّ بِهٖ كَثِيْرًا وَّيَهْدِيْ بِهٖ كَثِيْرًاۗ وَمَا يُضِلُّ بِهٖٓ اِلَّا الْفٰسِقِيْنَۙ

Artinya : "Sesungguhnya Allah tidak segan membuat perumpamaan seekor nyamuk atau yang lebih kecil dari itu. Adapun orang-orang yang beriman, mereka tahu bahwa itu kebenaran dari Tuhan. Tetapi mereka yang kafir berkata "Apa maksud Allah dengan perumpamaan ini ?". Dengan (perumpamaan) itu banyak orang yang dibiarkan-Nya sesat, dan dengan itu banyak (pula) orang yang diberi-Nya petunjuk. Tetapi tidak ada yang dia sesatkan dengan (perumpamaan) itu selain orang-orang fasik ".

Menurut Quraisy Shehab ayat di atas menjelaskan bahwa Allah tidak berkeberatan dengan penyebutan nyamuk dalam Al-Quran, meskipun nyamuk dianggap kecil, tidak berguna dan membawa virus penyakit. Allah SWT menurunkan ayat ini untuk memberitahukan kepada kaum musyrik bahwa Allah tidak segan-segan menciptakan binatang seperti nyamuk (ba'ūdhah) bahkan Allah pun tidak segan-segan menciptakan binatang yang lebih kecil dari nyamuk yaitu virus (fauqa ba'ūdhah).

Berdasarkan ayat di atas, berkaitan dengan penyakit yang sedang menimpa dunia yaitu penyakit covid-19, dimana Allah menciptakan binatang yang lebih kecil dari seekor nyamuk yaitu virus dan bahkan bisa menyebabkan meninggal dunia.

2.9.2 Pandangan Al-Qur'an Mengenai Estimasi

Sebagaimana firman Allah Swt, dalam Al-Qur'an surah Az-Zumar ayat 47 mengenai estimasi atau perkiraan yang berbunyi :

وَلَوْ اَنَّ لِلَّذِيْنَ ظَلَمُوْا مَا فِى الْاَرْضِ جَمِيْعًا وَّمِثْلَهٗ مَعَهٗ لَافْتَدَوْا بِهٖ مِنْ سُوْۤءِ الْعَذَابِ يَوْمَ الْقِيٰمَةِۗ وَبَدَا لَهُمْ مِّنَ اللّٰهِ مَا لَمْ يَكُوْنُوْا يَحْتَسِبُوْنَ

Artinya : “Dan sekiranya orang-orang yang dzalim mempunyai apa yang ada dibumi semuanya dan (ada pula) sebanyak itu besertanya, niscaya mereka akan menebus dirinya dengan itu dari siksa yang buruk pada hari kiamat. Dan jelaslah bagi mereka adzab dari Allah yang belum pernah mereka perkirakan”.

Pada ayat tersebut Allah menjelaskan bahwa seandainya orang-orang musyrikin yang dzalim itu mempunyai seluruh kekayaan yang ada dimuka bumi dan ditambah sebanyak itu pula, niscaya mereka akan menebus dirinya dengan kekayaan itu untuk azab yang buruk dan dahsyat yang akan ditimpakan kepada mereka pada hari Kiamat.

Berdasarkan penjelasan ayat di atas, bahwa ada kaitan ayat tersebut dengan metode estimasi atau perkiraan yang terletak pada ada potongan ayat “ يَحْتَسِبُوْنَ” yang artinya “Perkiraan”. Pada kata “Perkiraan” ditujukan pada orang-orang yang muslim dan dzalim tentang azab Allah yang tidak pernah terpikirkan sebelumnya.