BAB II

TINJAUAN PUSTAKA

2.1 Regresi Poisson

Regresi Poisson adalah analisis yang di gunakan untuk mengetahui hubungan antara variabel respon (Y) dan variabel prediktor (X) dimana varibel respon diasumsikan berdistribusi poisson. Distribusi poisson sering digunakan untuk kejadian – kejadian yang jarang terjadi (Trivedi, 1998).

2.1.1. Model Regresi Poisson

Regresi Poisson adalah model regresi yang dapat digunakan pada data yang variabel responnya berdistribusi tidak normal dan berjenis diskrit, yaitu berdistribusi Poisson sebagai syarat utamanya. Distribusi Poisson memberikan suatu model yang realistis untuk berbagai macam fenomena acak selama nilai dari variabel acak Poisson berupa bilangan bulat non-negatif (Wulandari S,dkk, 2010).

Fungsi distribusi Poisson dituliskan dengan persamaan :

$$p(y;\mu) = \frac{e^{-\mu}\mu^{y}}{y!} (y = 0,1,2, \dots.) \qquad (2.2)$$

Persamaan di atas digunakan untuk menghitung peluang variabel acak Y, di mana mean dan variansi distribusi Poisson adalah sama, yaitu E[Y] = μ_Y = var[Y]. Kondisi ini disebut juga dengan *equidispersion*. Karena nilai ekspektasi sama dengan variansi maka sebaran faktor akan berpengaruh terhadap lainnya, sehingga asumsi homogenitas tidak harus dipenuhi pada data Poisson (Khosghoftar & Szabo, 2004). Untuk mengetahui apakah data yang diamati berdistribusi Poisson atau tidak, dapat dilakukan dengan uji *Kolmogorov-smirnov*, dimana hipotesis pengujiannya sebagai berikut:

$H_0 : F(y) = F^*(y)$ (variabel random y mengikuti distribusi poisson)

$H_1 : F(y) = F^*(y)$ (variabel random y mengikuti distribusi poisson)

Kriteria pengujian dalam uji *Kolmogorov-smirnov* adalah tolak H_0 jika nilai signifikansi < α . Regresi Poisson pilihan yang tepat, ketika variabel respon Y merupakan bilangan bulat yang nilainya kecil. Model regresi Poisson ditulis sebagai berikut Myers, (1990) :

$$y_i = \mu_i + \varepsilon_i (i = 1,2, \dots n) \tag{2.4}$$

dimana y_i adalah jumlah kejadian dan μ_i adalah rata-rata jumlah kejadian di waktu tertentu, dan diasumsikan tidak berubah. Persamaan distribusi Poisson dinyatakan dengan persamaan sebagai berikut :

$$p(y_i; \hat{\beta}) = \frac{e^{-[\mu(x_i;\hat{\beta})]}[t_i \mu(x_i; \hat{\beta})]^{y_i}}{y_i!} \tag{2.5}$$

dimana $\mu(x_i; \widehat{\beta})$ adalah rata-rata Poisson dan vektor $\hat{\beta}$ menunjukkan taksiran dari parameter. Mean dan varians untuk model regresi Poisson adalah sebagai berikut :

$$\mu_i = t_i \mu_i(x_i; \hat{\beta}) = t_i \, exp(x_i^T \hat{\beta})$$

Dan

$Var(y_i) = t_i \mu_i(x_i; \hat{\beta}) = t_i \, exp(x_i^T \hat{\beta})$

Selanjutnya model regresi Poisson dapat ditulis sebagai berikut (Myers,1990) :

$$y = t_i \, exp(x_i^T \hat{\beta}) + \varepsilon_i \tag{2.6}$$

2.1.2. Penaksiran Parameter Model Regresi Poisson

Fungsi likelihood dari persamaan distribusi poisson (2.5) adalah sebagai berikut (Myers,1990) :

$$\begin{aligned} L(y, \hat{\beta}) &= \prod_{i=1}^{n} p\,(y_i; \hat{\beta}) \\ &= \prod_{i=1}^{n} \left\{ \frac{[t_i \mu(x_i; \hat{\beta})]^{y_i} e^{-[t_i \mu(x_i; \hat{\beta})]}}{y_i!} \right\} \\ &= \prod_{i=1}^{n} \frac{[t_i \mu(x_i; \hat{\beta})]^{y_i} e^{-\sum_{i=1}^{n} t_i \mu(x_i; \hat{\beta})}}{\prod_{i=1}^{n} y_i!} \end{aligned} \tag{2.7}$$

Untuk memaksimalkan persamaan diatas makan digunakan teknik iteratif yang dapat menghasilkan estimator maximum likelihood untuk koefisien regresi dalam $\hat{\beta}$. Prosedur yang disarankan oleh (Myers, 1990) untuk menemukan penaksir maximum likelihood adalah pendekatan *Iteratively Reweighted Least Squares (IRWLS).*

IRWLS menggunakan metode Nweton-Raphson, yang umumnya pada iterasi ke-s, metode Newton-Raphson memperbaiki taksiran $\hat{\beta}$, yang biasa dipakai dengan rumus (Cameron, A.C danTrivendi, P.K; 1998):

$$\hat{\beta}_{s+1} = \hat{\beta}_s - \hat{H}_s^{-1}\hat{g}_s \quad (2.8)$$

dimana $g = \frac{\partial lnL(y;\hat{\beta})}{\partial\hat{\beta}}$ dan $H = \frac{\partial^2 lnL(y;\hat{\beta})}{\partial(\hat{\beta})^2}$.

Metode Newton-Raphson yang digunakan untuk menyelesaikan persamaan berikut adalah :

$$\frac{\partial lnL(y;\hat{\beta})}{\partial\hat{\beta}} = 0 \quad (2.9)$$

dimana

$$\ln L(y;\hat{\beta}) = \sum_{i=1}^{n} y_i \ln t_i\, \mu(x_i;\hat{\beta}) - \sum_{i=1}^{n} t_i\, \mu(x_i;\hat{\beta}) - \sum_{i=1}^{n} \ln(y_i)\,! \quad (2.10)$$

Persamaan *likelihood* untuk mencari $\hat{\beta}$ adalah sebagai berikut :

$$\sum_{i=1}^{n} \frac{y_i}{\mu(x_i;\hat{\beta})} \frac{\partial\mu(x_i;\hat{\beta})}{\partial\hat{\beta}} - \sum_{i=1}^{n} t_i \frac{\partial\mu(x_i;\hat{\beta})}{\partial\hat{\beta}} = 0$$

$$\sum_{i=1}^{n} \left[\frac{y_i}{\mu(x_i;\hat{\beta})} - t_i\right]\left[\frac{\partial\mu(x_i;\hat{\beta})}{\partial\hat{\beta}}\right] = 0 \quad (2.11)$$

2.1.3. Pengujian Parameter Model Regresi Poisson

Parameter yang dihasilkan dari serangkaian proses penaksiran belum tentu memberi pengaruh yang signifikan terhadap variabel respon. Untuk menguji apakah parameter model memiliki pengaruh yang signifikan perlu dilakukan pengujian terhadap parameter model tersebut. Pengujian signifikasi parameter model regresi poisson terdiri dari uji serentak dan uji parsial.

a. Pengujian Serentak

Menggunakan metode *Maximum Likelihooh Ratio Test* (MLRT) dimana hipotesis pengujiannya adalah sebagai berikut :

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

H_1 : Paling sedikit $\beta_1 \neq 0$; j=1,2,…,p

Statistik uji yang digunakan adalah :

$$D(\hat{\beta}) = -2ln\Lambda$$

$$= -2ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right)$$

$$= 2\left(lnL(\hat{\Omega}) - lnL(\hat{\omega})\right) \quad (2.12)$$

Tolak H_0 jika nilai devians model regresi poisson atau *likelihood ratio* atau $D(\hat{\beta}) > X^2_{(\alpha,p)}$ artinya ada salah satu parameter yang berpengaruh secara signifikan terhadap variabel dependen. Λ merupakan rasio antara fungsi *likelihood* untuk himpunan parameter di $H_0\left(L(\hat{\omega})\right)$ dengan fungsi *likelihood* dengan himpunan parameter selain parameter di bawah $H_0\left(L(\hat{\Omega})\right)$.

b. Pengujian Individu

Pengujian signifikasi parameter secara parsial dengan hipotesis sebagai berikut:

H_0 : $\beta_j = 0$

H_0 : $\beta_j \neq 0$

Statistik Uji : $Z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$ (2.13)

Tolak H_0 jika $|Z_{hitung}| > Z_{\alpha/2}$ artinya parameter ke-j signifikanterhadap model regresi poisson. n adalah jumlah sampel, k adalah banyaknya variabel, dan α adalah taraf signifikasi.

2.2 Multikolinieritas

Sebelum melakukan analisis regresi poisson, hal pertama yang harus dilakukan adalah mendeteksi adanya multikolinieritas menurut (Hocking R, 1996) dapat dilihat dari nilai (*Variance Inflation Factors*) VIF. Persamaannya dapat di tulis :

$$VIF_j = \frac{1}{1 - R_j^2}$$

Dengan R_j^2 adalah nilai koefisien determinasi. nilai VIF lebih dari 10 menunjukkan variabel prediktor saling berkorelasi dengan variabel prediktor yang lain. Sedangkan nilai VIF kurang dari 10 menunjukkan tidak ada autokorelasi antar variabel prediktor. Apabila terjadi multikolinieritas solusi untuk mengatasi adanya Multikolinearitas adalah dengan mengeluarkan variabel prediktor yang tidak signifikan (*dropping variable*) dan meregresikan kembali variabel-variabel prediktor yang signifikan.

2.3 Overdispersi

Regresi Poisson memiliki asumsi equidispersi, yaitu kondisi dimana nilai rataan dan ragam pada peubah Y bernilai sama. Penyimpangan asumsi yang sering terjadi pada regresi Poisson yaitu overdispersi. Kondisi overdispersi yaitu nilai ragam lebih besar dari pada nilai rataan pada peubah Y. Salah satu akibatnya adalah simpangan baku dari penduga parameter menjadi berbias ke bawah dan signifikansi dari peubah penjelas menjadi berbias ke atas, sehingga menghasilkan kesimpulan yang tidak valid (Ismail dan Jemain 2007). Taksiran dispersi diukur dengan *devians* atau *pearson's Chi Square* yang dibagi derajat bebas. Data overdispersi jika taksiran dispersi lebih besar dari 1 dan underdispersi jika taksiran dispersi kurang dari 1, atau dengan *overdispersion test* dengan melihat p-value > nilai α. (Khoshgoftaar, 2004 dalam Camelia P.S, Nur I.M dan Darsyah M.Y, 2016)

2.4 Heterogenitas Spasial

Adanya salah satu aspek spasial yaitu sifat heterogenitas spasial adalah sebagai syarat bisa dilakukan pemodelan data dengan menggunakan pendekatan titik dengan GWPR. Untuk mendeteksi ada atau tidaknya heterogenitas spasial dalam model dilakukan uji *Breusch-pagan* dengan hipotesis sebagai berikut :

$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$

Untuk $i \neq j$, dengan $i, j = 1,2, \dots, n$.

Statistik uji : $BP = \frac{1}{2} h^T Z (Z^T Z)^{-1} Z^T h$ (2.14)

Dengan $h = \frac{e_i^2}{\sigma^2} - 1$, Z adalah vektor variabel respon y yang berukuran (nx1) dan telah dinormal standarkan. e_i^2 adalah kuadrat galat untuk pengamatan ke-i dan σ^2 adalah ragam dari e_i. Jika tolak H_0 ($BP < \alpha$) maka aspek spasial terpenuhi dan selanjutnya dilakukan analisis lebih lanjut dalam penelitian menggunakan pendekatan GWPR.

2.5 Model GWPR

GWPR adalah suatu metode pengembangan dari regresi poisson yang membedakan adalah dalam permodelan GWPR memperhatikan pembobot berupa letak lintang dan bujur dari titik-titik pengamatan yang diamati yang dinotasikan dengan (u_i, v_i) yang merupakan vektor koordinat dua dimensi lokasi ke-i. Model GWPR merupakan model regresi linier lokal yang menghasilkan penaksir parameter model yang bersifat lokal untuk setiap titik atau lokasi dimana data tersebut dikumpulkan (Nakaya *et al*,2004). Model GWPR dapat ditulis sebagai berikut :

$\ln(\mu_i) = \beta_0(u_i, v_i) + \sum_{j=i}^{p} \beta_j(u_i, v_i)x_{ij}$ atau

$$\mu_i = exp\left(\beta_0(u_i, v_i) + \sum_{j=i}^{p} \beta_j(u_i, v_i)x_{ij}\right) \quad (2.15)$$

Keterangan :

μ_i = nilai observasi variabel respon ke-i

x_{ij} = nilai observasi variabel prediktor j pada pengamatan ke-i

u_i = koordinat spasial *latitude* pengamatan ke-i

v_i = koordinat spasial *longitude* prngamatan ke-i

$\beta_0, \beta_1, \dots, \beta_p$ = koefisien regresi variabel X ke-j

2.5.1. Penaksiran Parameter Model GWPR

Penaksir parameter pada model GWPR sama seperti pada model regresi poisson yaitu menggunakan MLE dengan menambahkan faktor pembobot letak geografis pada fungsi *ln-likelihood*-nya sehingga didapatkan persamaan sebagai berikut:

$$\ln L^*\left(\beta(u_i,v_i)\right) = \sum_{i=1}^{n}\left(-exp\left(x_i^T\beta(u_i,v_i)\right) + y_i x_i^T\beta(u_i,v_i) - ln(y_i!)\right) w_{ij}(u_i,v_i) \quad (2.16)$$

Untuk memperoleh estimasi parameter yaitu dengan mendiferensiasikan persamaan (2.16) terhadap $\beta(u_i,v_i)$ dan hasilnya harus sama dengan nol.

$$\frac{\partial lnL^*\left(\beta(u_i,v_i)\right)}{\partial\beta^T(u_i,v_i)} = \sum_{i=1}^{n}\left[y_i x_i - x_i exp\left(x_i^T\beta(u_i,v_i)\right)\right] w_{ij}(u_i,v_i)$$

$$\sum_{i=1}^{n}\left[y_i x_i - x_i exp\left(x_i^T\beta(u_i,v_i)\right)\right] w_{ij}(u_i,v_i) = 0 \quad (2.17)$$

Karena persamaan (2.17) masih berbentuk implisit, maka digunakan iterasi numerik dengan metode Newton-Raphson. Iterasi Newton-Raphson adalah :

$$\beta_{(m+1)}(u_i,v_i) = \beta_m(u_i,v_i) - H_{(m)}^{-1}\left(\beta_m(u_i,v_i)\right) g_{(m)}\left(\beta_m(u_i,v_i)\right) \quad (2.18)$$

Dimana

$$g_{(m)}\left(\beta_m(u_i,v_i)\right) = \frac{\partial lnL^*\left(\beta_m(u_i,v_i)\right)}{\partial\beta^T(u_i,v_i)} \quad (2.19)$$

$$= -\sum_{i=1}^{n} x_i w_{ij}(u_i,v_i) exp\left(x_i^T\beta(u_i,v_i)\right) + \sum_{i=1}^{n} x_i w_{ij}(u_i,v_i) y_i$$

$$H_{(m)}\left(\beta_m(u_i,v_i)\right) = \frac{\partial lnL^{*2}\left(\beta_m(u_i,v_i)\right)}{\partial\beta(u_i,v_i)\partial\beta^T(u_i,v_i)}$$

$$= -\sum_{i=1}^{n} x_i w_{ij}(u_i,v_i) exp\left(x_i^T\beta(u_i,v_i)\right) \qquad (2.20)$$

Iterasi akan berhenti jika sudah didapatkan keadaan konvergen dimana

$\left\|\beta_{(m+1)}(u_i,v_i) - \beta_m(u_i,v_i)\right\| \leq \varepsilon$, nilai ε adalah 10^{-5} (Aulele,2010)

2.5.2. Pengujian Parameter Model GWPR

Pengujian kesamaan antara model regesi poisson dan GWPR dilakukan menggunakan perbandingan nilai devians model regresi poisson dan model GWPR. Hipotesisnya adalah sebagai berikut:

$H_0 : \left(\beta_j(u_i,v_i)\right) = \beta_j$

$H_1 : \left(\beta_j(u_i,v_i)\right) \neq \beta_j; i = 1,2, \ldots, n; j = 0,1,2, \ldots, p$

Statistiak Uji : $F_{hitung} = \frac{Devians\ Model\ ^A/_{dfA}}{Devians\ Model\ ^B/_{dfB}}$ (2.22)

Keterangan : dfA = derajat bebas untuk regresi poisson

dfB = derajat bebas untuk GWPR

Kriteria pengujiannya adalah tolak H0 jika $F_{hitung} > F_{(\alpha,dfA,dfB)}$artinya ada perbedaan yang signifikan antara model regresi poisson dengan model GWPR. Pengujian selanjutnya yaitu pengujian parameter secara parsial untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel respon. Hipotesis untuk pengujian parameter secara parsial adalah:

$H_0 : \left(\beta_j(u_i,v_i)\right) = 0$

$H_1 : \left(\beta_j(u_i, v_i)\right) \neq 0$; i = 1,2,...,n; j = 0,1,2,...,p

Statistik uji : $t = \frac{\hat{\beta}_j(u_i,v_i)}{se\left(\hat{\beta}_j(u_i,v_i)\right)}$ (2.23)

Tolak H_0 jika $|t_{hitung}| > t_{(n-k-1;\alpha/2)}$ artinya parameter ke-j pada lokasi ke-i (u_i, v_i) berpengaruh signifikan terhadap model.

2.6 Penentuan Bandwith dan Pembobot

Bandwidth (*h*) merupakan radius suatu lingkaran dimana titik yang berada dalam radius lingkaran masih dianggap berpengaruh dalam membentuk model lokasi *i*. Menurut Fortheringham *et al*. (2002) salah satu metode yang bisa digunakan untuk menentukan bandwidth optimum adalah metode *Cross Validation* (CV) yang didefinisikan sebagai berikut:

$$CV = \sum_{i=1}^{n}\left(y_i - \hat{y}_{\neq i}(h)\right)^2 \qquad (2.24)$$

Dimana $\hat{y}_{\neq i}(h)$ merupakan nilai penaksir (*fitted value*) untuk y_i dengan radius *h*, tetapi pengamatan di titik *i* dikeluarkan dari proses penaksiran. *bandwidth* yang optimum diperoleh dari hasil nilai CV yang paling kecil.

Pada regresi global tanpa pembobotan geografis, masing-masing observasi memiliki nilai pembobotan sebesar 1, $w_{ij} = 1$; j=1,2,…,n dengan *j* menunjukan titik pada daerah dimana data diobservasi dan *i* menunjukan sebuah titik pada ruang untuk penaksiran parameter. Pembobotan bervariasi sesuai lokasi pada titik regresi ke-*i*, dimana $0 \leq w_{ij} \leq 1$ dan w_{ij} semakin kecil ketika jarak bertambah. Dengan kata lain observasi yang dekat dengan titik regresi akan memberikan bobot yang besar dibandingkan observasi yang jauh dari titik regresi. Setelah ditentukan nilai *bandwidth* yang optimum, selanjutnya adalah mencari nilai pembobot fungsi kernel dengan rumus sebagai berikut.

1. Adaptive Gaussian Kernel

   $$W_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right] \quad (2.25)$$

2. Adaptive Bisquare Kernel

   $$W_{ij} = \begin{cases} \left[1-\left(\frac{d_{ij}}{h_i}\right)^2\right]^2, untuk\ d_{ij} \le h_i \\ 0, untuk\ d_{ij} > h_i \end{cases} \quad (2.26)$$

   Dimana w_{ij}adalah nilai pembobot fungsi kernel dari lokasi pengamatan *j* untuk mengestimasi koefisien pada lokasi *i dan* d_{ij} merupakan jarak euclidien antara lokasi *i* ditulis (u_i, v_i) ke lokasi *j* dutulis (u_i, v_i), *h* adalah nilai bandwidth optimum. Untuk mencari jarak euclidean dicari dengan rumus berikut :

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

2.7 *Akaike Information Criterion* (AIC)

Indikator pemilihan model terbaik dari beberapa model yang umum digunakan pada pemodelan spasial adalah nilai *Akaike Information Criterion* (AIC). Perhitungan AIC dapat dilakukan dengan rumus:

$$AIC = D(h) + 2K(h) \quad (2.28)$$

D adalah nilai devians model dengan *bandwidth* (*h*) dan K adalah jumlah parameter dalam model dengan *bandwidth* (*h*). Untuk kriteria pemilihan model terbaik adalah model dengan AIC terkecil.

2.8 Angka Kematian Ibu

Pengertian Angka Kematian Ibu (*Maternal Mortality Rate*) adalah Jumlah kematian ibu akibat dari proses kehamilan, persalinan dan paska persalinan per 100.000 kelahiran hidup pada masa tertentu. Angka pengukuran risiko kematian wanita yang berkaitan dengan peristiwa kehamilan.

Kematian ibu adalah kematian wanita dalam masa kehamilan, persalinan dan dalam masa 42 hari (6 minggu) setelah berakhirnya kehamilan tanpa memandang usia kehamilan maupun tempat melekatnya janin, oleh sebab apa pun yang berkaitan dengan atau diperberat oleh kehamilan atau pengelolaannya, bukan akibat kecelakaan.

Kematian ibu dikelompokkan menjadi dua (2), yaitu

1. Kematian sebagai akibat langsung kasus kebidanan
2. Kematian sebagai akibat tidak langsung yaitu kasus kebidanan yang disebabkan penyakit yang sudah ada sebelumnya, atau penyakit yang timbul selama kehamilan dan bukan akibat langsung kasus kebidanan, tetapi diperberat oleh pengaruh fisiologi kehamilan.

Kematian wanita hamil akibat kecelakaan (misalnya kecelakaan mobil) tidak digolongkan sebagai kematian ibu.

Angka Kematian Ibu (AKI) atau *Maternal Mortality Rate* (MMR) berguna untuk menggambarkan tingkat kesadaran perilaku hidup sehat, status gizi dan kesehatan ibu, kondisi lingkungan, tingkat pelayanan kesehatan terutama untuk ibu hamil, pelayanan kesehatan waktu melahirkan dan masa nifas.

Beberapa determinan penting yang mempengaruhi AKI secara langsung antara lain status gizi, anemia pada kehamilan. Faktor mendasar penyebab kematian ibu maternal adalah tingkat pendidikan ibu, kesehatan lingkungan fisik maupun budaya, ekonomi keluarga, pola kerja rumah tangga.

Menteri Kesehatan Nafsiah Mboi mengatakan, terdapat dua penyebab ibu meninggal saat meninggal yakni infeksi dan perdarahan. Untuk yang penyebabnya infeksi sudah dapat ditekan karena sebagian besar kelahiran dilakukan di pusat layanan kesehatan seperti puskesmas, rumah sakit, klinik dan sebagainya.

Sementara untuk perdarahan disebabkan empat hal yakni :

1. Melahirkan ketika usia muda,
2. Melahirkan ketika usia tua,
3. Melahirkan terlalu sering dan,
4. Jarak antara satu kelahiran dan lainnya terlalu rapat.