BAB II

TINJAUAN PUSTAKA

2.1 Kolinieritas

Langkah awal sebelum melakukan analisis Regresi Poisson yaitu pendeteksian adanya kasus kolinieritas menurut (Hocking R, 1996) dapat dilihat nilai VIF (*Variance Inflation Factors*) > 10

Variance Inflation Factors (VIF) dinyatakan dalam :

$$VIF_j = \frac{1}{1-R_j^2} \tag{2.1}$$

R_j^2 menunjukkan nilai koefisien determinasi. Nilai VIF_j yang bernilai 1 menunjukkan bahwa variabel-variabel independent tidak saling berkorelasi, jika nilainya lebih dari 10 menunjukkan adanya kolinieritas antara variabel-variabel independen.

2.2 Model regresi Poisson

Model regresi Poisson merupakan model standar untuk data diskrit dan termasuk dalam model regresi nonlinier. Metode regresi poisson biasanya diterapkan pada penelitian kesehatan masyarakat, biologi dan teknik dimana variabel responnya (y) berupa cacahan objek yang merupakan fungsi dari sejumlah karakteristik tertentu (Wulandari S, dkk, 2010). Model regresi poisson ditulis sebagai berikut (Myers, 1990) :

http://repository.unimus.ac.id

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

dimana μ adalah mean distribusi Poisson. Parameter μ sangat bergantung pada beberapa unit yang ditetapkan atau periode waktu, jarak, luas, volume, dan lain-lain. Distribusi Poisson digunakan untuk memodelkan peristiwa yang jarang terjadi selama periode tertentu. Probabilitas banyak kejadian y dalam periode waktu t yaitu :

$$p(y;\mu) = \frac{e^{-\mu t}(\mu t)^{y}}{y!} \quad (y = 0,1,2, \dots.) \tag{2.3}$$

Persamaan tersebut digunakan untuk probabilitas kejadian y, dan rata-rata jumlah kejadian, berdasarkan asumsi bahwa rata-rata jumlah kejadian per periode waktu adalah konstan. Pengujian kesesuaian distribusi untuk variabel y adalah dengan uji Kolmogorov-Smirnov. Berikut ini adalah hipotesis uji Kolmogorov-Smirnov :

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

$H_1 : F(y) \neq F^*(y)$ (variabel random y tidak mengikuti distribusi tertentu)

Daerah penolakan untuk pengujian ini adalah tolak H_0 jika α pada taraf signifikansi $T > w_{1-\alpha/2}$, dimana $w_{1-\alpha/2}$ adalah nilai kuantil dari statistik uji Kolmogorov-Smirnov pada uji dua sisi.

Misal, apabila terdapat sekumpulan data dengan struktur sebagai berikut :

http://repository.unimus.ac.id

$$\begin{bmatrix} y_1 & x_{11} \cdots & x_{k1} \\ \vdots & \vdots & \vdots \\ y_n & x_{1n} \cdots & x_{kn} \end{bmatrix}$$

maka, model regresi Poissonnya dapat 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 dalam periode t_i . μ_i diasumsikan tidak berubah dari data ke data. 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 parameter yang ditaksir. 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) :

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

http://repository.unimus.ac.id

2.2.1 Estimasi Parameter Model Regresi Poisson

Berdasarkan persamaan distribusi Poisson yang ditujukkan pada persamaan (2.5), maka fungsi likelihoodnya adalah sebagai berikut (Myers,1990) :

$$L(y,\hat{\beta}) = \prod_{i=1}^{n} p\,(y_i;\hat{\beta})$$

$$= \prod_{i=1}^{n} \left\{ \frac{[t_i\mu(x_i;\hat{\beta})]^{y_i} e^{-[t_i\mu(x_i;\hat{\beta})]}}{y_i!} \right\}$$

$$= \prod_{i=1}^{n} \frac{[t_i\mu(x_i;\hat{\beta})]^{y_i} e^{-\sum_{i=1}^{n} t_i\mu(x_i;\hat{\beta})}}{\prod_{i=1}^{n} y_i!} \quad (2.7)$$

Persamaan diatas dimaksimalkan dengan menggunakan teknik iteratif yang menghasilkan penaksir 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 dan Trivendi, 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 $H = \frac{\partial^2 lnL(y;\hat{\beta})}{\partial(\hat{\beta})^2}$.

http://repository.unimus.ac.id

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

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

dimana

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

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

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

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

2.2.2 Pengujian Parameter Model Regresi Poisson

Pengujian kesesuaian model dengan *goodness of fit* disebut devians (Kleinbaum, *et. al*, 1988). Berikut ini adalah hipotesis pengujian kesesuaian model regresi Poisson.

$$H_0 : \mu_i = t_i\, \mu(x_i;\, \hat{\beta}), i = 1,2, \dots n$$

$$H_1 : \mu_i \neq t_i\, \mu(x_i;\, \hat{\beta})$$

Statistik uji yang digunakan adalah sebagai berikut:

$$G = -2l\left[\frac{L(y;\hat{\beta})}{L(y;\mu)}\right] \tag{2.12}$$

http://repository.unimus.ac.id

Nilai $\mu_i = t_i\mu(x_i;\beta)$, sehingga Persamaan (2.9) dapat ditulis menjadi persamaan berikut :

$$\ln L(y;\hat{\beta}) = \sum_{i=1}^{n} y_i \ln \mu_i - \sum_{i=1}^{n} \mu_i - \sum_{i=1}^{n} \ln(y_i)!$$

$$= \sum_{i=1}^{n} y_i \ln \hat{y}_i - \sum_{i=1}^{n} \hat{y}_i - \sum_{i=1}^{n} \ln(y_i)! \qquad (2.13)$$

Sedangkan nilai $L(y;\mu)$ dapat ditulis dalam persamaan berikut :

$$L(y,\mu) = \prod_{i=1}^{n} \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} = \frac{\left(\prod_{i=1}^{n} \mu_i^{y_i}\right) exp\left(-\sum_{i=1}^{n} \mu_i\right)}{\prod_{i=1}^{n} y_i!} \qquad (2.14)$$

Nilai $\hat{\mu}_i = y_i$, sehingga Persamaan (2.14) dapat juga ditulis menjadi sebagai berikut :

$$L(y,\hat{\mu}) = \frac{\left(\prod_{i=1}^{n} y_i^{y_i}\right) exp\left(-\sum_{i=1}^{n} y_i\right)}{\prod_{i=1}^{n} y_i!}$$

Nilai $\ln(y;\hat{\mu})$ dapat ditulis dalam persamaan berikut :

$$lnL(y;\hat{\mu}) = \sum_{i=1}^{n} y_i \ln y_i - \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \ln(y_i)! \qquad (2.15)$$

Persamaan (2.12) dapat ditulis menjadi berikut :

$$G = -2ln\left[L(y;\hat{\beta}) - L(y,\hat{\mu})\right]$$

$$= -2ln\left[L(y;\hat{\mu}) - L(y,\hat{\beta})\right]$$

$$= -\sum_{i=1}^{n}[(y_i \ln y_i - y_i - ln(y_i)!) - (y_i \ln \hat{y}_i - \hat{y}_i - ln(y_i)!)]$$

$$G = 2\sum_{i=1}^{n}\left[y_i ln\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i)\right] \quad (2.16)$$

Nilai G disebut devians untuk model regresi Poisson. Model yang sesuai devians mendekati distribusi *Chi-Square* dengan derajat bebas $(n - k - 1)$, dimana n adalah jumlah pengamatan dan $k + 1$ adalah jumlah parameter (Ismail N, Jemain A.A, 2005).

Devians seperti *Sum Square Error* pada analisis regresi berganda yaitu nilai data pengamatan sama dengan prediksi $(y_i = \hat{y}_i)$ maka nilai $G = 0$ (Kleinbaum D G, *et.al*, 2014). Besarnya nilai devians terjadi jika semakin besar selisih antara respon pengamatan dan respon taksiran, maka nilai devians juga akan semakin besar. Taksiran diharapkan dapat mendekati pengamatan atau tingkat kesalahan diharapkan bernilai kecil, sehingga nilai devians akan bernilai kecil, dan nilai devians yang diharapkan adalah nilai devians yang kecil.

Parameter model regresi Poisson yang telah dihasilkan dari estimasi parameter belum tentu mempunyai pengaruh yang signifikan terhadap model, maka perlu dilakukan pengujian terhadap parameter model regresi Poisson secara individu.

2.3 Overdispersi

Metode regresi Poisson mewajibkan equidispersi, yaitu kondisi dimana nilai mean dan varians dari variabel respon bernilai sama. Namun, adakalanya terjadi fenomena overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi berarti varians lebih besar daripada mean. Taksiran dispersi

http://repository.unimus.ac.id

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 (Khoshgoftaar dalam Camelia P.S, *et.al*, 2016).

2.4 Model *Zero-Inflated Poisson Regression*

Salah satu penyebab terjadinya overdisersi adalah lebih banyak observasi yang bernilai nol daripada yang ditaksir untuk model regresi Poisson. Salah satu metode analisis yang diusulkan untuk kasus seperti ini adalah model regresi *Zero-Inflated Poisson (ZIP)* (Jansakul, N. & Hinde, J. P dalam Nur, I.M, 2018).

Jika y_i adalah variabel random independen yang mempunyai distribusi ZIP, observasi diduga muncul dalam dua cara yang sesuai untuk state yang terpisah. State pertama terjadi dengan probabilitas ω_i dan menghasilkan hanya observasi bernilai nol, sementara state kedua terjadi dengan probabilitas $(1-\omega_i)$ dan berdistribusi Poissondengan mean μ_i. Proses dua state ini memberikan distribusi campuran dua komponen dengan fungdi probabilitas sebagai berikut (Jansakul, N. dan Hinde, J. P., 2002) :

$$Pr(Y_i = y_i) = \begin{Bmatrix} \omega_i + (1-\omega_i)e^{-\mu_i}, y_i = 0 \\ (1-\omega_i)\frac{e^{-\mu_i}\mu_i^{\mu_i}}{y_i!}, y_i = 1,2, \ldots 0 \leq \omega_i \leq 1 \end{Bmatrix} \quad (2.17)$$

Lambert dalam (Jansakul, N. dan Hinde, J. P., 2002) menunjukkan model golongan untuk μ dan ω sebagai berikut :

$$log(\mu) = X\beta \text{ dan } logit\ (\omega) = log\left(\frac{\omega}{1-\omega}\right) = X\gamma \quad (2.18)$$

http://repository.unimus.ac.id

Dimana X adalah matriks variabel prediktor sedangkan β dan γ adalah parameter yang akan ditaksir.

2.4.1 Estimasi Parameter Model Regresi *Zero-Inflated Poisson*

Estimasi parameter regresi ZIP dengan menggunakan metode Maximum Likelihood Estimation (MLE). Fungsi Likelihood ZIP adalah sebagai berikut (Khoshgoftaar, *et.al*, 2004) :

$$L(\beta,\gamma|y_i,x_i) = \left\{ \begin{array}{l} \prod_{i=1}^{n} \frac{exp(x_i^T\gamma)+exp\left(-exp(x_i^T\beta)\right)}{1+exp(x_i^T\gamma)}, y_i = 0 \\ \prod_{i=1}^{n} \frac{\frac{1}{1+exp\left(x_i^T(\gamma)\right)}\left(exp\left((-exp(x_i^T\beta)+(x_i^T\beta)y_i)\right)\right)}{y_i!}, y_i > 0 \end{array} \right\} \quad (2.19)$$

dimana x_i adalah variabel prediktor, y_i adalah variabel respon, serta β dan γ adalah parameter yang akan ditaksir.

Fungsi Log-Likelihood gabungan untuk model regresi ZIP diberikan oleh :

$$lnL(\beta,\gamma|y_i,x_i) = \sum_{\substack{i=1 \\ y_i=0}}^{n} ln\left(exp(x_i'\gamma)\right) - \sum_{i=1}^{n} ln\left(1+exp(x_i'\gamma)\right) + \sum_{\substack{i=1 \\ y_i>0}}^{n} \left((y_i x_i^T\beta) - exp(x_i^T\beta)\right) - \sum_{\substack{i=1 \\ y_{i>0}}}^{n} ln(y_i)\,! \quad (2.20)$$

Estimasi maximum likelihood untuk β dan γ dapat di peroleh dengan menggunakan pendekatan standart untuk model campuran, yaitu Algoritma EM.

http://repository.unimus.ac.id

Algoritma EM memberikan prosedur sederhana yang dapat diimplementasi dalam software standar, atau metode estimasi langsung seperti metode Newton-Raphson.

2.4.2 Pengujian Peremeter Model Zero-Inflated Poisson

Pengujian kesesuaian model regresi ZIP adalah dengan menggunakan *Likelihood Ratio (LR) test.* Hipotesis untuk pengujian kesesuaian model adalah sebagai berikut :

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

$$H_1: paling\ sedikit\ ada\ satu\ \beta_r \neq 0\ atau\ \gamma_r \neq 0, 0 < r < k-1$$

dimana k+1 adalah jumlah parameter, β_r adalah parameter model log ke-r, dan γ_r adalah parameter model logit ke-r. Lestari (2008) telah melakukan perhitungan statistik uji untuk pengujian kesesuaian model sebagai berikut :

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

$$= \left(2\sum_{i=1}^{n}\left(z_i x_i^T\hat{\gamma} - ln\left(1 + exp(x_i^T\hat{\gamma})\right) + 2\sum_{i=1}^{n}(1-z_i)\left(y_i x_i^T\hat{\beta} - exp\left(x_i^T\hat{\beta}\right)\right)\right) - 2\sum_{i=1}^{n} z_i\hat{\gamma}_0 - ln(1 + x_i^T\hat{\gamma}_0) + 2\sum_{i=1}^{n}(1 - z_i)\left(y_i\hat{\beta}_0 - exp\left(\hat{\beta}_0\right)\right)\right) \quad (2.21)$$

Sedangkan pengujian parameter secara individu ada dua, yaitu pengujian parameter model log dan pengujian parameter model logit. Statistik uji untuk

http://repository.unimus.ac.id

pengujian parameter model log secara individu adalah sebagai berikut (Lestari,2008) :

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

$$= \left(2\sum_{i=1}^{n}\left(z_i x_i^T\hat{y} - ln\left(1 + exp(x_i^T\hat{\gamma})\right) + 2\sum_{i=1}^{n}(1-z_i)\left(y_i x_i^T\hat{\beta} - exp(x_i^T\hat{\beta})\right)\right) - 2\sum_{i=1}^{n}(1-z_i)\left(y_i x_i^T\hat{\beta}_i - exp(x_i^T\hat{\beta}_i)\right)\right) \quad (2.22)$$

Statistik uji yang digunakan untuk pengujian parameter model logit adalah sebagai berikut (Lestari,2008) :

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

$$= 2\sum_{i=1}^{n}\left(z_i x_i^T\hat{y} - ln\left(1 + exp(x_i^T\hat{\gamma})\right) + 2\sum_{i=1}^{n}(1-z_i)\left(y_i x_i^T\hat{\beta} - exp(x_i^T\hat{\beta})\right)\right) - 2\sum_{i=1}^{n}(1-z_i)ln(y_i)! - 2\sum_{i=1}^{n}\left(z_i\hat{\gamma}_0 - ln\left(1 + exp(\hat{\gamma}_0)\right)\right) \quad (2.23)$$

Daerah penolakan untuk ketiga pengujian diatas adalah tolak H_0 pada taraf signifikansi α jika $G_{hitung} > \mathcal{X}^2_{(v,a)}$, dimana v adalah derajat bebas.

2.4.3 Pemilihan Model Terbaik

Metode Akaike's Information Criterion (AIC) merupakan salah satu metode yang digunakan untik memilih model terbaik untuk regresi ZIP. Nilai AIC adalah sebagai berikut (Dalrymple,*et.al.*,2001) :

$$AIC = G + (k + 1) \qquad (2.24)$$

dimana G adalah statistik uji kesesuaian model, dan $k + 1$ adalah jumlah parameter. Model terbaik regresi ZIP adalah model dengan nilai AIC terkecil.

2.5 Tetanus Neonatorum

2.5.1 Definisi Tetanus Neonatorum

Tetanus adalah penyakit dengan tanda utama kekakuan otot (spasme) tanpa disertai gangguan kesadaran yang disebabkan oleh kuman *Clostridium tetani*. Gejala ini bukan disebabkan kuman secara langsung, tetapi sebagai dampak eksotoksin (tetanospasmin) yang dihasilkan oleh kuman pada sinaps ganglion sambungan sumsum tulang belakang, sambungan neuromuskular (*neuromuscular junction*) dan saraf otonom (Pusponegoro dkk, 2004).

2.5.2 Faktor Penyebab Tetanus Neonatorum

Terjadinya tetanus neonatorum karena disebabkan oleh beberapa faktor diantaranya adalah

a. Cakupan imunisasi TT2+ terhadap ibu hamil

   Sebagai upaya mengendalian infeksi tetanus neonatorum yang merupakan salah satu faktor resiko kematian ibu dan kematian bayi, maka dilaksanakan program imunasasi Tetanus Toksoid (TT) bagi Wanita Usia Subur (WUS) dan ibu hamil. Cakupan imunisasi TT2+ merupakan imunisasi yang didapatkan dari TT2 sampai TT5.

http://repository.unimus.ac.id

b. Presentase tenaga bidan terhadap jumlah bayi

Bagi ibu hamil yang berada pada daerah dengan akses sulit, kebijakan kementrian kesehatan yaitu mengembangkan program Kemitraan Bidan dan Dukun serta Rumah Tunggu Kelahiran. Oleh karena itu, Kementerian Kesehatan tetap konsisten dalam menerapkan kebijakan bahwa seluruh persalinan harus ditolong oleh tenaga kesehatan dan didorong dilakukan di fasilitas pelayanan kesehatan. Kebijakan Dana Alokasi Khusus (DAK) Bidang Kesehatan menggariskan bahwa pembangunan puskesmas harus satu paket dengan rumah dinas tenaga kesehatan. Dengan disediakan rumah tinggal, maka tenaga kesehatan termasuk bidan akan siaga di tempat tugasnya dan dapat memberikan pertolongan persalinan setiap saat. Dan penanganan bayi pun dapat dilaksanakan dengan baik dengan adanya bidan.

c. Presentase ibu bersalin ditolong tenaga kesehatan

Pertolongan persalinan oleh tenaga kesehatan sangat penting, karena dapat menurunkan resiko kematian ibu dan bayi yang disebabkan oleh penyakit tertentu serta penggunaan alat yang streil untuk menguranginya. Namun demikian, meskipun persalinan ditolong oleh tenaga kesehatan tetapi tidak dilaksanakan di fasilitas pelayanan kesehatan, dianggap menjadi salah satu penyebab masih tingginya angka kematian ibu.

http://repository.unimus.ac.id

d. Presentase cakupan kunjungan neonates lengkap terhadap jumlah bayi

Kunjungan neonatal pertama (KN1) adalah cakupan pelayanan kesehatan bayi baru lahir (umur 6 jam-48 jam) di satu wilayah kerja pada kurun waktu tertentu yang ditangani sesuai standar oleh tenaga kesehatan terlatih di seluruh sarana pelayanan kesehatan. Selain KN1, indikator yang menggambarkan pelayanan kesehatan bagi neonatal adalah Kunjungan Neonatal Lengkap (KN lengkap) yang mengharuskan agar setiap bayi baru lahir memperoleh pelayanan Kunjungan Neonatal minimal tiga kali sesuai standar di satu wilayah kerja pada kurun waktu satu tahun.

e. Presentase kunjungan ibu hamil K4

Penilaian terhadap pelaksanaan pelayanan kesehatan ibu hamil dapat dilakukan dengan melihat cakupan K1 dan K4. Cakupan K1 adalah jumlah ibu hamil yang telah memperoleh pelayanan antenatal pertama kali oleh tenaga kesehatan dibandingkan jumlah sasaran ibu hamil di satu wilayah kerja pada kurun waktu satu tahun. Sedangkan cakupan K4 adalah jumlah ibu hamil yang telah memperoleh pelayanan antenatal sesuai dengan standar paling sedikit empat kali sesuai jadwal yang dianjurkan di tiap trimester dibandingkan jumlah sasaran ibu hamil di satu wilayah kerja pada kurun waktu satu tahun.

http://repository.unimus.ac.id