Jurnal Diferensial, Volume 04, Nomor 02, November 2022
E-ISSN: 2775-9644

PENGGUNAAN MODEL REGRESI QUASI LIKELIHOOD UNTUK MENGATASI MASALAH OVERDISPERSI PADA REGRESI POISSON

Robertus D. Guntur[1*], Keristina Br. Ginting[1*], Ganesha L. Putra[1], Jeanete Y. Nenabu[1],
1. Program Studi Matematika FST UNDANA

*Penulis korespondensi: robertus_guntur@staf.undana.ac.id

ABSTRAK

Salah satu masalah utama dalam menganalisis data menggunakan model regresi poisson adalah overdispersi. Kehadiran kondisi ini dalam model menyebabkan estimasi koefisien regresi yang diperoleh cenderung menyimpang dari nilai parameter sebenarnya. Tujuan utama dari penelitian ini adalah untuk menunjukkan bagaimana menyelesaikan masalah overdispersi dengan menggunakan model regresi quasi likelihood. Penelitian ini menggunakan data sekunder dengan ukuran sampel sebanyak 316. Hasil penelitian ini menunjukkan bahwa nilai khi kuadrat pearson dan nilai Devians dalam pengujian Goodness of Fit pada regresi quasi likelihood mendekati nilai distribusi khi kuadrat dengan derajat bebas 313. Dengan demikian penggunaan model regresi Quasi Likelihood lebih memadai dari pada model regresi Poisson dalam menangani data yang teroverdispersi.

Kata Kunci : Regresi poisson, overdispersi, regresi quasi likelihood

1. PENDAHULUAN

Pada umumnya salah satu metode peramalan yang sering digunakan adalah analisis regresi. Analisis regresi mempelajari hubungan fungsional antara variabel respon dan variabel bebas yang dinyatakan dalam sebuah persamaan matematis. Variabel yang mudah didapat atau tersedia digolongkan ke dalam variabel bebas (variabel prediktor), sedangkan variabel yang terjadi karena variabel bebas itu merupakan variabel tak bebas (variabel respon). Untuk selanjutnya, variabel bebas dinyatakan dengan $X_1, X_2, \ldots, X_k$ ($k \geq 1$), sedangkan variabel respon dinyatakan dengan Y [1]–[3].

Dalam analisis regresi, akan ditentukan hubungan fungsional yang diharapkan berlaku untuk populasi berdasarkan data sampel yang diambil dari populasi yang bersangkutan. Hubungan fungsional ini akan dituliskan dalam bentuk persamaan matematis yang selanjutnya disebut sebagai persamaan regresi. Analisis regresi yang umumnya digunakan adalah analisis regresi klasik, dimana variabel responnya merupakan data kontinu yang mengikuti distribusi normal. Namun dalam perkembangannya, model regresi klasik ini tidak mampu mengatasi permasalahan-

permasalahan dimana variabel respon berupa data diskrit dan tidak berdistribusi normal.

Solusi untuk mengatasi masalah tersebut adalah dengan menggunakan model linier tergenalisir atau biasa disebut *Generalized Linear Model* (GLM). Uji asumsi yang diterapkan pada GLM tidak mengharuskan asumsi kenormalan dari variabel respon dan juga tidak mengharuskan kehomogenan dari variansinya [4]. Salah satu model regresi yang termasuk dalam penerapan GLM adalah regresi *Poisson*.

Regresi *Poisson* digunakan untuk menganalisis hubungan antara variabel respon yang berupa data diskrit dengan satu atau lebih variabel indepeden, dimana nilai mean dan variansinya diasumsikan sama (equidispersi). Pada prakteknya seringkali data cacah memperlihatkan variasi yang sangat besar dimana mean sampel lebih besar daripada variansi sampel (underdispersi) atau sebaliknya variansi sampel lebih besar dari pada mean sampel (overdispersi). Jika model Poisson standar tetap diterapkan pada data yang mengandung overdispersi, nilai estimasi parameternya menjadi kurang efisien. Karena overdispersi akan menghasilkan penduga atau taksiran koefisien regresi yang cenderung menyimpang dari nilai parameter sebenarnya. Galat baku yang dihasilkan juga ternyata *underestimate*. Artinya galat baku lebih kecil dari nilai sesungguhnya dan sebagai konsekuensinya model tersebut akan memberikan kesimpulan yang keliru dalam interprestasinya [3].

Salah satu metode yang dapat digunakan untuk mengatasi masalah overdispersi pada regresi Poisson adalah dengan pendekatan model regresi *quasi likelihood*. Dalam model regresi *quasi likelihood* diasumsikan $Var(y_i) = \phi\mu_i$, dimana ϕ merupakan parameter dispersi. Berdasarkan uraian diatas maka penulis tertarik untuk mengkaji tentang model regresi *Quasi Likelihood* dalam mengatasi masalah overdispersi pada regresi *Poisson*.

2. METODE

2.1 Sumber data

Penelitian ini dilakukan dengan menggunakan data sekunder yang diperoleh dari http://www.ats.ucla.edu/stat/sas/dae/poissonreg.csv. Besarnya sampel sebanyak 316 siswa. Variabel responnya adalah jumlah hari ketidakhadiran siswa, sementara itu nilai ujian Matematika dan Seni ditetapkan sebagai variabel prediktornya.

2.2 Metode Analisis

Adapun langkah-langkah yang ditempuh dalam penelitian ini adalah sebagai berikut:

1. **Pemodelan awal dengan menggunakan model regresi Poisson.**

Model regresi Poisson diturunkan dari distribusi Poisson dengan memparameterisasi hubungan antara parameter mean dan variabel prediktor **X** [5]–[7] . Sehigga persamaan distribusi Poisson dapat dinyatakan sebagai

$$P(Y;\boldsymbol{\beta}) = \frac{\left(\mu(\mathbf{X}_i,\boldsymbol{\beta})\right)^Y e^{-\mu(\mathbf{X}_i,\boldsymbol{\beta})}}{Y!}$$

Bentuk $\mu(\mathbf{X}_i,\boldsymbol{\beta})$ diartikan sebagai mean μ dianggap sebagai sebuah fungsi dari variabel-variabel prediktor $x_1, x_2, ..., x_N$ dan $\boldsymbol{\beta}$ merupakan nilai koefisien regresi yang kemudian ditaksir dan menyatakan seberapa besar pengaruh prediktor terhadap variabel respon.

$\boldsymbol{\beta}$ diestimasi dengan menggunakan metode kuadrat terkecil biasa, sehingga persamaan Likelihoodnya adalah

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} = \sum_i (y_i - \mu_i) x_{ij} = 0 \quad ; \quad j = 1, 2, ...$$

2. **Pengujian terjadinya over dispersi pada model regresi Poisson yang dihasilkan.**

Dalam rangka hendak melihat apakah dalam pemodelan dengan model *Poisson* terdapat overdipersi atau tidak dalam pengamatan yang diperoleh, hipotesis berikut diujikan [5], [8]:

$H_0 : \text{var}(y) = \mu$ versus $H_1 : \text{var}(y) = \phi\mu$

Statistik ujinya diberikan oleh :

$$W = \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2 - y_i}{\sqrt{2\sum_{i=1}^{N} \hat{\mu}_i^{\,2}}}$$

W berdistribusi Khi-Kuadrat dengan derajat bebasnya 1. H_0 ditolak jika $W > \chi^2_{0.05}(1)$.

3. **Pengujian Kecocokan Model *(Goodness of Fit) pada regresi Poisson***

Pengujian kecocokan model di uji dengan menggunakan statistik Khi-Kuadrat Pearson dan nilai devians [6]. Statistik Khi-Kuadrat Pearson dinyatakan sebagai berikut :

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2}{\text{var}(y_i)}.$$

Nilai devians diperoleh dari

$$D = 2L(y;y) - L(\hat{\mu};y)$$

dengan $L(\hat{\mu};y)$ dan $L(y;y)$ merupakan nilai log-likelihood model yang dievaluasi dengan nilai $\hat{\mu}$ dan *y*. Untuk model yang cukup atau memadai, *D* secara asimptotik akan sama dengan distribusi Khi-Kuadrat dengan derajat bebas *N – p*. Oleh karena itu, jika nilai Khi-Kuadrat Pearson dan Devians model hampir sama dengan derajat bebasnya maka model itu dianggap memadai [7].

4. **Pemodelan dengan menggunakan model regresi *Quasi Likelihood***

Untuk N observasi independen dengan parameter ϕ tidak diketahui, diperoleh *quasi likelihood* sebagai berikut

$$Q_n\left(y;\mu,\phi\right)=\sum_{i=1}^{n}Q\left(y_i;\mu_i,\phi\right)$$

Estimasi parameter β dalam model Regresi *Quasi Likelihood* diperoleh dari :

$$\frac{\partial Q_i}{\partial \beta_j}=\sum_{i=1}^{N}\frac{\left(y_i-\mu_i\right)}{\phi}x_{ij}=0$$

5. **Pengujian Kecocokan Model *(Goodness of Fit)* pada** *Quasi Likelihood*

Pengujian kecocokan model pada regresi *Quasi Likelihood* juga di uji dengan menggunakan statistik Khi-Kuadrat Pearson dan nilai devians seperti yang diuraikan dalam tahapan ketiga.

6. **Membandingkan model yang dihasilkan.**

Pada tahap ini dilakukan perbandingan antara nilai Khi Kuadrat Pearson dan devians yang diperoleh dari model regresi Poisson and model regresi *Quasi Likelihood.* Suatu model dikatakan memadai jika nilai Khi Kuadrat Pearson dan devians mendekati nilai distribusi Khi Kuadrat table dengan derajat bebas yang sesuai dengan data [6], [9], [10]

3. HASIL DAN PEMBAHASAN

3.1 Pemodelan dengan Regresi *Poisson*

Pemodelan dalam regresi poisson diawali dengan mengestimasi parameter model yang mana estimasi awal diperoleh dengan menggunakan metode regresi linear biasa pada data siswa sebanyak 316, dimana

$$\mathbf{X}=\begin{bmatrix}1 & 56,99 & 42,45\\ 1 & 37,09 & 46,82\\ \vdots & \vdots & \vdots\\ 1 & 6,75 & 13,13\end{bmatrix}\text{ dan }\mathbf{Y}=\begin{bmatrix}4\\ 4\\ \vdots\\ 12\end{bmatrix}$$

sehingga diperoleh :

$$\hat{\boldsymbol{\beta}}_{(0)} = \left[\mathbf{X'X}\right]^{-1}\left[\mathbf{X'Y}\right]$$

$$= \begin{bmatrix} 316 & 15405.31962 & 15820.15885 \\ 15405.31962 & 851737.168 & 840845.3234 \\ 15820.15885 & 840845.3234 & 893388.9825 \end{bmatrix}^{-1} \begin{bmatrix} 1836 \\ 82680.05 \\ 84256.54 \end{bmatrix}$$

$$= \begin{bmatrix} 10.01924 \\ -0.02961 \\ -0.05524 \end{bmatrix}$$

Dengan demikian $\hat{\mu}_{i(1)} = \exp\left(10.01924 - 0.02961x_{1i} - 0.05524x_{2i}\right)$.

Estimasi selanjutnya adalah

$$\hat{\boldsymbol{\beta}}_{(1)} = \left[\mathbf{X'WX}\right]_{(0)}^{-1}\left[\mathbf{X'WY}\right]_{(0)}$$

$$= \begin{bmatrix} 9.014071997 \\ -0.029446932 \\ -0.05494032 \end{bmatrix}$$

Iterasi terus berlanjut hingga konvergen yaitu jika $\left|\dfrac{\hat{\boldsymbol{\beta}}_{(r+1)} - \hat{\boldsymbol{\beta}}_{(r)}}{\hat{\boldsymbol{\beta}}_{(r)}}\right| \leq 10^{-6}$. Taksiran parameter regresi *Poisson* diperoleh pada iterasi ke 11 dan hasilnya disajikan dalam Tabel 3.1.

Tabel 3.1. Hasil iterasi untuk koefisien regresi *Poisson* data nilai Matematika, Bahasa dan Seni, dan jumlah ketidakhadiran siswa

Iterasi	Parameter		
	β_0	β_1	β_2
0	10.0192	-0.0296	-0.0552
1	9.01407	-0.0294	-0.0549
2	8.00025	-0.029	-0.0541
3	6.96421	-0.0278	-0.0521
4	5.87691	-0.025	-0.0471
5	4.70461	-0.0192	-0.0369
6	3.4634	-0.0179	-0.0141
7	2.69848	-0.008	-0.0108
8	2.45261	-0.0051	-0.0094
9	2.43447	-0.0049	-0.0093
10	2.43438	-0.0049	-0.0093
11	2.43438	-0.0049	-0.0093

sehingga dengan model *Poisson*, data ini dapat direpresentasikan dengan persamaan

$$\ln\left(\hat{\mu}_i\right) = 2.434381 - 0.0049x_{1i} - 0.00927x_{2i}.$$

Estimasi standar error

Setelah proses iterasi konvergen, galat baku setiap koefisien dapat diperoleh dengan memperhatikan matriks **I.**

$$\mathbf{I}^{11} = \mathbf{X'W}^{11}\mathbf{X}$$

$$= \begin{bmatrix} 1836.000001 & 82680.05474 & 84256.54232 \\ 82680.05474 & 4333965.7 & 4231354.683 \\ 84256.54232 & 4231354.683 & 4487750.473 \end{bmatrix}$$

Invers matriks ini adalah

$$\left[\mathbf{I}^{11}\right]^{-1} = \left[\mathbf{X'W}^{11}\mathbf{X}\right]^{-1}$$

$$= \begin{bmatrix} 0.004470909 & -4.20307E-05 & -4.4311E-05 \\ -4.20307E-05 & 3.29908E-06 & -2.3215E-06 \\ -4.4311E-05 & -2.3215E-06 & 3.2436E-06 \end{bmatrix}$$

Sehingga nilai standar error dari parameter model regresi poisson adalah

$$SE\left(\beta_0\right) = \sqrt{0.004470909} = 0.066864855$$

$$SE\left(\beta_1\right) = \sqrt{3.29908\text{E}-06} = 0.001816337$$

$$SE\left(\beta_2\right) = \sqrt{3.2436\text{E}-06} = 0.001801001$$

Pengujian Signifikansi Model

Hipotesis yang digunakan adalah

$$H_0 : \beta_j = 0 \quad \text{melawan} \quad H_1 : \beta_j \neq 0 \quad j = 0,1,...$$

dengan statistik uji Wald-Kuadrat, yaitu $Z^2 = \left(\frac{\hat{\beta}_j}{SE\left(\hat{\beta}_j\right)}\right)^2$.

Statistik Wald Kuadrat yang diperoleh untuk β_0, β_1 dan β_2 secara berturut-turut adalah 1325.505, 7.791 dan 26.465. Karena $\chi^2_{0.05}(1) = 3.84$ maka semua parameter β berpengaruh secara signifikan dalam model regresi Poisson.

3.2 Pengujian overdispersi

Hipotesis yang digunakan adalah

$$H_0 : \text{var}\left(y\right) = \mu \quad \text{versus} \quad H_1 : \text{var}\left(y\right) = \phi\mu.$$

Dengan statistik uji $W = \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2 - y_i}{\sqrt{2\sum_{i=1}^{N} \hat{\mu}_i^{\,2}}}$.

Dari hasil perhitungan diperoleh nilai $W = 99.6$, sementara $\chi^2_{0,05}(1) = 3,84.$ Oleh karena itu H_0 ditolak, dengan kata lain data mengalami overdispersi.

3.3 Pengujian Kecocokan Model *(Goodness of Fit) pada regresi Poisson*

Pengujian kecocokan model pada regresi Poisson dilakukan dengan menghitung nilai Khi Kuadrat Pearson dan nilai devians seperti diuraikan pada bagian berikut.
Nilai Khi Kuadrat Pearson untuk model *Poisson* adalah

$$\sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2}{var(y_i)} = \sum_{i=1}^{316} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} = 2996.433$$

Nilai devians untuk model *Poisson* diperoleh dari :

$$D = 2\left[L(y;y) - L(\hat{\mu};y)\right]$$
$$= 2\left[\left(\sum_i \left(y_i \ln(y_i) - y_i - \ln(y_i!)\right)\right) - \left(\sum_i \left(y_i \ln(\hat{\mu}_i) - \hat{\mu}_i - \ln(y_i!)\right)\right)\right] = 2\left[\sum_i y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - y_i + \hat{\mu}_i\right]$$

Maka nilai devians untuk model *Poisson* adalah

$$D = 2\left[\sum_{i=1}^{316} y_i\, ln\left(\frac{y_i}{\hat{\mu}_i}\right) - y_i + \hat{\mu}_i\right]$$
$$= 2[1.15\text{E} + 03]$$
$$= 2304.302$$

3.4 Pemodelan dengan Regresi *Quasi Likelihood*

Sebagaimana ditunjukkan dalam bagian 3.2 bahwa data terindikasi overdispersi, sehingga pemodelan dengan model *Quasi Likelihood* dapat dilaksanakan. Estimasi awal diperoleh dengan menerapkan metode regresi linear biasa. Seperti pada pemodelan dengan regresi Poisson diperoleh $\hat{\beta}_{0(0)} = 10.01924,$ $\hat{\beta}_{1(0)} = -0.02961$ dan $\hat{\beta}_{2(0)} = -0.05524$, sehingga $\hat{\mu}_{i(1)} = \exp(10.01924 - 0.02961x_{1i} - 0.05524x_{2i})$ dan dengan mengambil nilai $\phi = 1$, maka $w_{i(1)} = \hat{\mu}_{i(1)}$. Iterasi selanjutnya menjadi

$$\hat{\boldsymbol{\beta}}_{(1)} = \left[\mathbf{X'WX}\right]_{(0)}^{-1}\left[\mathbf{X'WY}\right]_{(0)}$$

$$= \begin{bmatrix} \sum_i^N w_{i(1)} & \sum_i^N w_{i(1)}x_{1i} & \sum_i^N w_{i(1)}x_{2i} \\ \sum_i^N w_{i(1)}x_{1i} & \sum_i^N w_{i(1)}x_{1i}^{\ 2} & \sum_i^N w_{i(1)}x_{1i}x_{2i} \\ \sum_i^N w_{i(1)}x_{2i} & \sum_i^N w_{i(1)}x_{1i}x_{2i} & \sum_i^N w_{i(1)}x_{2i}^{\ 2} \end{bmatrix}_{(0)}^{-1} \times \begin{bmatrix} \sum_{i=1}^N w_{i(1)}\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}\right) \\ \sum_{i=1}^N w_{i(1)}x_{1i}\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_2}\right) \\ \sum_{i=1}^N w_{i(1)}x_{2i}\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_N}\right) \end{bmatrix}_{(0)}$$

$$= \begin{bmatrix} \sum_i^N \hat{\mu}_i & \sum_i^N \hat{\mu}_i x_{1i} & \sum_i^N \hat{\mu}_i x_{2i} \\ \sum_i^N \hat{\mu}_i x_{1i} & \sum_i^N \hat{\mu}_i x_{1i}^{\ 2} & \sum_i^N \hat{\mu}_i x_{1i}x_{2i} \\ \sum_i^N \hat{\mu}_i x_{2i} & \sum_i^N \hat{\mu}_i x_{1i}x_{2i} & \sum_i^N \hat{\mu}_i x_{2i}^{\ 2} \end{bmatrix}_{(0)}^{-1} \begin{bmatrix} \sum_{i=1}^N \hat{\mu}_i\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}\right) \\ \sum_{i=1}^N \hat{\mu}_i x_{1i}\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_2}\right) \\ \sum_{i=1}^N \hat{\mu}_i x_{2i}\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_N}\right) \end{bmatrix}_{(0)}$$

$$= \begin{bmatrix} 1.01116\text{E}-05 & -1.81102\text{E}-07 & -1.03923\text{E}-07 \\ -1.81102\text{E}-07 & 1.95823\text{E}-08 & -1.32976\text{E}-08 \\ -1.03923\text{E}-07 & -1.32976\text{E}-08 & 1.88507\text{E}-08 \end{bmatrix}^{-1} \begin{bmatrix} 2194646.452 \\ 48043425.4 \\ 43075029.69 \end{bmatrix}$$

$$= \begin{bmatrix} 9.014071997 \\ -0.029446932 \\ -0.05494032 \end{bmatrix}$$

Dengan kata lain, estimasi parameter regresi pada regresi *Quasi Likelihood* menghasikan nilai yang sama seperti pada estimasi parameter regresi *Poisson*. Hasil iterasi untuk regresi *Quasi Likelihood* dapat dilihat pada tabel 3.2.

Tabel 3.2. Hasil iterasi untuk koefisien regresi *Quasi Likelihood* data nilai Matematika, Bahasa dan Seni, dan jumlah ketidakhadiran siswa

Iterasi	Parameter		
	β_0	β_1	β_2
0	10.0192	-0.0296	-0.0552
1	9.01407	-0.0294	-0.0549
2	8.00025	-0.029	-0.0541
3	6.96421	-0.0278	-0.0521
4	5.87691	-0.025	-0.0471

5	4.70461	-0.0192	-0.0369
6	3.4634	-0.0179	-0.0141
7	2.69848	-0.008	-0.0108
8	2.45261	-0.0051	-0.0094
9	2.43447	-0.0049	-0.0093
10	2.43438	-0.0049	-0.0093
11	2.43438	-0.0049	-0.0093

Sehingga dengan model *Quasi Likelihood*, data ini dapat direpresentasikan dengan persamaan

$$\ln\left(\hat{\mu}_i\right) = 2.434381 - 0.0049x_{1i} - 0.00927x_{2i}.$$

Estimasi standar error

Galat baku setiap koefisien dapat diperoleh dengan memperhatikan matriks **I,** yaitu

$$\mathbf{I}^{11} = \mathbf{X'W}^{11}\mathbf{X}$$

$$= \begin{bmatrix} 1836.000001 & 82680.05474 & 84256.54232 \\ 82680.05474 & 4333965.7 & 4231354.683 \\ 84256.54232 & 4231354.683 & 44877750.473 \end{bmatrix}$$

Invers matriks ini adalah

$$\left[\mathbf{I}^{11}\right]^{-1} = \left[\mathbf{X'W}^{11}\mathbf{X}\right]^{-1}$$

$$= \begin{bmatrix} 0.004470909 & -4.20307E-05 & -4.4311E-05 \\ -4.20307E-05 & 3.29908E-06 & -2.3215E-06 \\ -4.4311E-05 & -2.3215E-06 & 3.2436E-06 \end{bmatrix}$$

Untuk regresi *Quasi Likelihood*, galat baku setiap koefisien diperoleh dari mengalikan standar eror asal dengan $\sqrt{\frac{\chi^2}{(N-p)}}$.

Sehingga diperoleh

$$SE\left(\beta_0\right) = \left(\sqrt{0.004470909}\right)\left(\sqrt{\frac{2996.432736}{(316-3)}}\right)$$

$$= (0.066864855)(9.573267527)$$

$$= 0.640115143$$

$$SE(\beta_1) = \left(\sqrt{3.29908\text{E}-06}\right)\left(\sqrt{\frac{2996.432736}{(316-3)}}\right)$$
$$= (0.001816337)(9.573267527)$$
$$= 0.017388279$$
$$SE(\beta_2) = \left(\sqrt{3.2436\text{E}-06}\right)\left(\sqrt{\frac{2996.432736}{(316-3)}}\right)$$
$$= (0.001801001)(9.573267527)$$
$$= 0.017241463$$

Karena nilai standar erornya lebih besar, maka dapat dikatakan model ini adalah model yang lebih baik.

Pengujian signifikansi model

Hipotesis yang digunakan adalah

$$H_0 : \beta_j = 0 \quad \text{melawan} \quad H_1 : \beta_j \neq 0 \quad j = 0,1,\ldots$$

dengan statistik uji Wald, yaitu $Z^2 = \left(\frac{\hat{\beta}_j}{SE\left(\hat{\beta}_j\right)}\right)^2$.

Statistik Wald yang diperoleh untuk β_0, β_1 dan β_2 secara berturut-turut adalah 14.463 dan 0.079serta 0,289. Karena $\chi^2_{0.05}(1) = 3.84$ maka hanya β_0 yang berpengaruh secara signifikan dalam model regresi Quasi Likelihood.

3.5 Pengujian Kecocokan Model *(Goodness of Fit) pada* regresi Quasi Likelihood.

Pengujian kecocokan model pada regresi Quasi Likelihood dilakukan dengan menghitung nilai Khi Kuadrat Pearson dan nilai devians seperti diuraikan pada bagian berikut.

Perhitungan Khi Kuadrat Pearson dan nilai devians untuk model Quasi Likelihood adalah sebagai berikut :

Nilai Khi Kuadrat Pearson untuk model Quasi Likelihood adalah

$$\sum_{i=1}^{N} \frac{\left(y_i - \hat{\mu}_i\right)^2}{\text{var}\left(y_i\right)} = \sum_{i=1}^{316} \frac{\left(y_i - \hat{\mu}_i\right)^2}{\phi\hat{\mu}_i} = 313$$

Sedangkan nilai devians yang diberikan oleh model regresi *Quasi Likelihood* adalah

$$
\begin{aligned}
D &= 2\left[Q(y_i;y_i) - Q(\hat{\mu}_i;y_i)\right] \\
&= 2\left[\left(\sum_{i=1}^{N} y_i \ln y_i - y\right) - \left(\sum_{i=1}^{N} y_i \ln \hat{\mu}_i - \hat{\mu}_i\right)\right] \\
&= 2\left[\sum_{i=1}^{N} y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - y_i + \hat{\mu}_i\right]
\end{aligned}
$$

Dengan demikian, devians untuk model *Quasi Likelihood* diperoleh

$$
\begin{aligned}
D &= 2\sum_{i=1}^{316}\left[y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - y_i + \hat{\mu}_i\right] \\
&= 2 \times 3.13\mathrm{E} + 02 \\
&= 626
\end{aligned}
$$

3.6 Perbandingan Model *Quasi Likelihood* terhadap model *Poisson*

Studi ini menunjukkan bahwa model *Poisson* dan model *Quasi Likelihood* memberikan nilai yang sama untuk koefisien-koefisien regresinya, tetapi tidak untuk nilai standar errornya. Standar error koefisien regresi *Poisson* lebih kecil daripada yang diberikan oleh model *Quasi Likelihood*. Hasil penelitian ini konsisten dengan penelitian sebelumnya [11], [12] yang menunjukkan bahwa penerapan model quasi likelihood dalam regresi poisson mampu memperbaiki nilai standar error dari parameter model, dengan menghasilkan nilai standar error parameter yang lebih besar dari regresi poisson. Lebih lanjut penelitian ini menunjukkan bahwa nilai Khi Kuadrat Pearson dan devians yang diberikan oleh model regresi Poisson adalah 2996.433 dan 2304.302. Sedangkan untuk model regresi *Quasi Likelihood* memberikan Khi Kuadrat Pearson dan devians sebesar 313 dan 626. Sementara itu jika diambil $\alpha = 5\%$, maka ${\chi^2}_{313} = 355.26$. Karena nilai Khi Kuadrat Pearson dan devians dari model regresi *Quasi Likelihood* yang lebih mendekati nilai ${\chi^2}_{313}$ maka dapat dikatakan bahwa model regresi *Quasi Likelihood* lebih memadai dari pada model regresi *Poisson* dalam menangani data yang teroverdispersi.

4. SIMPULAN

Model regresi *Quasi Likelihood* digunakan sebagai model alternatif untuk menangani data yang terindikasi overdispersi. Dalam penelitian ini, nilai Khi Kuadrat Pearson dan nilai Devians dalam pengujian Goodness of Fit pada regresi *Quasi Likelihood* mendekati nilai distribusi Khi Kuadrat dengan derajat bebas 313. Dengan demikian dapat dikatakan bahwa model regresi *Quasi Likelihood* lebih memadai dari pada model regresi *Poisson* dalam menangani data yang teroverdispersi.

DAFTAR PUSTAKA

[1] R. B. Darlington and A. F. Hayes, *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*. Guilford Publications, 2016. [Online]. Available: https://books.google.co.id/books?id=YDgoDAAAQBAJ

[2] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*. Wiley, 2015. [Online]. Available: https://books.google.co.id/books?id=27kOCgAAQBAJ

[3] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*. Wiley-Interscience, 2003. [Online]. Available: https://books.google.co.id/books?id=IdSwtAEACAAJ

[4] I. M. Tirta, *Model Statistika Linier*. Jember: Universitas Jember, 2006.

[5] R. Winkelmann, *Econometric Analysis of Count Data*. Springer Berlin Heidelberg, 2013. [Online]. Available: https://books.google.co.id/books?id=BRjpCAAAQBAJ

[6] J. M. Hilbe, *Modeling Count Data*. Cambridge University Press, 2014. [Online]. Available: https://books.google.co.id/books?id=aZLfAwAAQBAJ

[7] R. Winkelmann, *Count Data Models: Econometric Theory and an Application to Labor Mobility*. Springer Berlin Heidelberg, 2013. [Online]. Available: https://books.google.co.id/books?id=-IjnCAAAQBAJ

[8] R. Zainordin, “Count Data Analysis using Poisson Regression and Handling of Overdispersion. [http://eprints.utm.my/12417/1/RaihanaZainordinMFS2009.pdf]. Faculty of Science, Universiti Teknologi Malaysia,” Faculty of Science, Universiti Teknologi Malaysia, Malaysia, 2009. [Online]. Available: http://eprints.utm.my/12417/1/RaihanaZainordinMFS2009.pdf

[9] J. M. Hilbe, *Negative Binomial Regression*. Cambridge University Press, 2011. [Online]. Available: https://books.google.co.id/books?id=DDxEGQuqkJoC

[10] A. Agresti, *An Introduction to Categorical Data Analysis*. Wiley, 2018. [Online]. Available: https://books.google.co.id/books?id=pHZyDwAAQBAJ

[11] R. A. Puspita, “Model Regresi Quasi-Likelihood Untuk Mengatasi Masalah Overdispersi Pada Regresi Poisson.,” Universitas Diponegoro, Semarang, 2011.

[12] D. P. Prami Meitriani, K. G. Sukarsa, And I. P. E. N. Kencana, “Penerapan Regresi Quasi-Likelihood Pada Data Cacah (Count Data) Yang Mengalami Overdispersi Dalam Regresi Poisson,” *E-J. Mat. Vol 2 No 2 2013DO - 1024843MTK2013v02i02p036*, May 2013, [Online]. Available: https://ojs.unud.ac.id/index.php/mtk/article/view/6290