Prosiding Seminar Nasional Penelitian, Pendidikan dan Penerapan MIPA
Fakultas MIPA, Universitas Negeri Yogyakarta, 16 Mei 2009

Maximum Likelihood Estimation Model Linear dan Log-Linear dalam Regresi Poisson

Yuli Wibawati, Jaka Nugraha

Jurusan Statistika FMIPA-UII
Kampus Terpadu UII, Jln Kaliurang KM 14. Telp/fax : 0274-895920

Abstrak
Regresi Poisson digunakan untuk pemodelan pada respon berdistribusi Poisson. Parameter rata-rata merupakan fungsi dari variabel independen. Telah dilakukan estimasi terhadap parameter dalam regresi Poisson menggunakan metode *Maximum Likelihood Estimation* (MLE). Estimasi didasarkan pada model linear dan model log-linear. Kedua model dibandingkan menggunakan data simulasi. Karena parameter rata-rata diasumsikan bernilai positif, hal ini mengakibatkan estimasi parameter dalam model log linear lebih sesuai dibanding model linear.
.
Kata Kunci: Maximum Likelihood, log-linear, hessian.

PENDAHULUAN

Analisis regresi merupakan metode statistika yang populer digunakan untuk menyatakan hubungan antara variabel dependen dengan variabel independen. Dalam menganalisis hubungan antara beberapa variabel, terdapat sejumlah fenomena dimana variabel dependennya berbentuk biner ataupun berbentuk diskrit. McCullagh dan Nelder (1989) mengamati suatu fenomena dimana variabel dependennya berbentuk diskrit tapi tidak biner, maka fenomena ini menyangkut banyaknya suatu kejadian dalam distribusi Poisson atau merupakan proses kejadian Poisson, dimana batas atas banyaknya kejadian itu tidak terbatas.

Di bawah kondisi eksperimental yang ideal jika terjadi suatu peristiwa sukses yang saling bebas dan dalam satu waktu atau ruang yang sama, maka model poisson merupakan suatu model pendekatan untuk banyaknya suatu kejadian yang diamati (Hajarisman, 2001). Selanjutnya, bilamana dihadapankan pada variabel dependen yang tidak kontinu, maka ada beberapa masalah yang akan dihadapi dan perlu mendapati perhatian khusus. Proses kejadian Poisson merupakan pengembangan dari proses Bernoulli dan binomial, dapat dinyatakan model regresi dari *n* buah data.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_k + \varepsilon_i \quad (1)$$

dengan (i = 1, 2, …, n dan y_i = 0, 1)

Jika diasumsikan bahwa $E(\varepsilon_i) = 0$, maka:

$$E(y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_k \quad (2)$$

Sekarang, jika $E(y_i) = P_i$ adalah proporsi populasi dari pengamatan pada x_{1i}, x_{2i}, ..., x_{ki}, untuk y = 1, atau dengan kata lain:

$P_i = \text{Prob}[y_i = 1]$
$Q_i = 1 - P_i = \text{Prob}[y_i = 0]$
dengan (i = 1, 2, . . ., n)

Jadi untuk n buah pengamatan akan terdapat n peluang P_1, P_2, …, P_n yang masing – masing merupakan parameter dari distribusi Bernoulli. Jika diperhatikan pada model di atas, jelas bahwa ε_i tidak kontinu jika hanya dua harga yang mungkin yaitu:

$$\varepsilon_i = y_i - [\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_k] \quad (3)$$

dengan harga bisa $1 - P_i = Q_i$ atau $0 - P_i = -P_i$. Akibatnya tidak ada asumsi tentang kenormalan pada model galatnya. Asumsi kedua yang pasti dilanggar adalah tentang kehomogenan variansi. Jika $E(\varepsilon_i) = 0$, maka:

$$\begin{aligned} \mathrm{Var}(\varepsilon_i) &= E(\varepsilon_i)^2 \\ &= (Q_i)^2 \,\mathrm{Prob}[y_i = 1] + (-P_i)^2 \,\mathrm{Prob}[y_i = 0] \\ &= Q^2 P_i + P_i^2 Q_i \\ &= P_i Q_i (Q_i + P_i) = P_i Q_i \end{aligned}$$

Jika P_i harganya bervariasi menurut tingkat variabel independen, maka varians galatnya, menjadi tidak homogen. Jadi, dalam hal ini ada dua pelanggaran asumsi, yaitu:

- Distribusi dari ε_i adalah diskrit dan tidak normal.
- Varians galat tidak homogen.

Dengan demikian penggunaan motode kuadrat terkecil biasa akan sulit dilakukan untuk kasus seperti yang dijelaskan di atas (Myers, 1990).

Hajarisman (2002), dalam penelitiannya melakukan pengembangan pada model regresi mengenai hubungan antara variabel dependen dengan variabel independen, dimana variabel dependennya merupakan bentuk diskrit yang tidak biner. Banyaknya suatu kejadian dalam suatu unit tertentu diasumsikan sebagai fungsi dari satu atau lebih variabel independen. Dalam penelitiannya peneliti mengasumsikan bahwa rata – rata dari banyaknya kejadian dalam unit tertentu merupakan parameter dari distribusi Poisson. Rata – rata Poisson ini juga merupakan fungsi dari variabel independen yang dapat dikembangkan menjadi model regresi Poisson, dimana fungsi *link* yang digunakan dalam pembentukan model regresi Poisson adalah $\hat{\lambda}_i = e^{x_i'\beta}$.

Studi kasus yang digunakan pada penelitian ini merupakan data simulasi dengan menggunakan perangkat lunak (*software*) R yaitu untuk membangkitkan data yang mengikuti distribusi Poisson. Metode yang digunakan untuk mengestimasi parameter dalam regresi Poisson adalah *Maximum Likelihood Estimation* (MLE). Estimasi didasarkan pada model linear dan model log linear, dengan mengasumsikan bahwa parameter rata – rata bernilai positif.

Regresi Normal

Ratih (Yasin, 2003) mendefinisikan Regresi normal merupakan model linear yaitu $Y_i = X_i\beta + \varepsilon_i$, untuk $i = 1, 2, 3, \ldots, N$. Pada awalnya dibangun untuk mempelajari hubungan antara variabel independen (X_i) dengan variabel dependen (Y_i) dimana Y_i berdistribusi normal dan saling bebas. Model linear yang paling awal dan paling sederhana ini biasa disebut sebagai model linear normal/klasik (*Normal/Classical Linear Model*, untuk selanjutnya disingkat NLM). Beberapa ciri dari bentuk linear ini adalah sebagai berikut ini.

1. Hubungan antara ekspektasi $E(Y) = \mu$ dan kombinasi (variabel independen) linier $\eta_i = x_i\beta$ adalah bersifat langsung (identitas), yaitu $\mu_i = \eta_i = x_i\beta$.
2. Bentuk varians dari variabel dependen yang bersifat konstan, yaitu varians $Y_i = \mathrm{var}(Y_i) = \sigma^2$. Dengan demikian, secara keseluruhan, $\mathrm{var}(Y) = \mathbf{I}_N\sigma^2$, dengan $\mathbf{I}_N$ adalah matriks identitas berordo N.

Pendugaan parameter untuk model NLM ini diturunkan dari persamaan turunan pertama dari fungsi log-likelihoodnya, yang setelah melalui beberapa penyederhanaan, dalam bentuk matriks dapat dinyatakan dengan:

$$S(\beta) = \left(\frac{\partial\mu}{\partial\beta}\right) V(Y)^{-1}(Y - \mu) = 0 \tag{4}$$

Hubungan antara rataan (μ) $\mu_i = x_i\beta$ dalam NLM, menghasilkan bentuk khusus

$$S(\beta) = \frac{1}{\sigma^2} X^T (Y - \mu) = 0 \tag{5}$$

Dalam kenyataannya, di lapangan sering dijumpai data yang tidak berdistribusi normal. Apabila data yang tidak berdistribusi normal ini masih saling bebas, maka model linear yang mempelajari hubungan variabel jenis data ini disebut model linear tergeneralisasi (*Generalized Linear Models,* untuk selanjutnya disingkat GLM).

Regresi Poisson

GLM secara detail oleh McCullagh dan Nelder (1989). Pendekatan ini didasarkan pada model regresi dengan distribusi responnya termasuk dalam kelurga eksponensial. Fungsi distribusi yang diakomodasi oleh GLM adalah distribusi yang termasuk dalam keluarga ekponensial, seperti distibusi binomial, normal, Poisson, gamma, eksponensial (Myers dan Montgomery, 1997).

Ada beberapa komponen utama dalam analisis GLM yaitu dimana variabel dependen memiliki sebaran yang termasuk dalam keluarga eksponensial (misalnya normal, binomial, Poisson, gamma) dan terdapat hubungan antara mean populasi dan prediktor yang dinyatakan oleh fungsi *link*. Adanya hubungan yang sesungguhnya tidak linear antara variabel dependen dan variabel independen, sehingga fungsi *link* dapat dianggap sebagai transformasi yang melinearkan hubungan tersebut.

Model untuk regresi Poisson pada dasarnya menyatakan rata – rata dari distribusi yang diskrit (dalam hal ini adalah proses Poisson) sebagai fungsi dari variabel independennya, berikut merupakan fungsi dari distribusi Poisson.

$$f(y;\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!} \text{, untuk } y = 0, 1, 2, .. \tag{6}$$

Estimasi Parameter

Penggunaan model regresi Poisson terdapat beberapa pelanggaran asumsi mengenai galat yang tidak berdistribusi normal dan variansi galat yang tidak homogen, sehingga dalam penaksiran parameter tidak bisa menggunakan metode kuadrat terkecil biasa. Untuk mengatasi hal tersebut maka metode yang dapat digunakan untuk mengestimasi parameter yaitu dengan menggunakan metode maksimum likelihood (Myers, 1990). Rata – rata dalam regresi Poisson dimodelkan sebagai fungsi dari sejumlah variabel independen. Adapun untuk mengestimasi parameter yaitu dengan menggunakan fungsi likelihood dan persamaan likelihood yang didasari dari distribusi Poisson

$$L(y,\lambda) = \prod_{i=1}^{n} f(y_i,\lambda) \tag{7}$$

$$L(y,\lambda) = \prod_{i=1}^{n}\left\{\frac{\lambda^{y_i}e^{-\lambda}}{y_i!}\right\} = \frac{\left\{\prod_{i=1}^{n}\lambda^{\sum_{i=1}^{n}y_i}\right\}e^{-\sum_{i=1}^{n}\lambda}}{\prod_{i=1}^{n}y_i!} \tag{8}$$

Pendugaan parameter dalam regresi Poisson dengan menggunakan *Maximum Likelihood Estmation* (MLE) dapat didekatkan pada model linear dan model log linear, yaitu:

- **Estimasi Parameter dengan Pendekatan Model Linear**

Model untuk regresi Poisson pada dasarnya menyatakan rata – rata dari distribusi yang diskrit sebagai fungsi dari variabel independennya. Berikut merupakan model regresi Poisson dengan pendekatan model linear:

$$\hat{\lambda} = \hat{\beta}_0 + \sum_{j=1}^{m}\hat{\beta}_j x_{ij} \tag{9}$$

Pendugaan parameter dengan menggunakan pendekatan model linear dengan metode MLE merupakan turunan pertama dari fungsi log likelihoodnya, maka bentuk yang akan digunakan untuk memperoleh parameter dengan menggunakan metode MLE adalah:

Turunan pertama:

$$\frac{\partial \ln L(y,\beta)}{\partial\beta} = 0 \tag{10}$$

sehingga dipunyai:

$$L\left(y,\beta\right)=\sum_{i=1}^{n} y_i \ln\left(\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}\right)-\sum_{i=1}^{n}\left(\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}\right)-\sum_{i=1}^{n}\ln\left(y_i!\right) \qquad (11)$$

dan dihasilkan:

$$\sum_{i=1}^{n}\left(\left(\frac{y_i}{\left(\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}\right)}-\left(\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}\right)\right)\left(\frac{\partial\left(\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}\right)}{\partial\hat{\beta}}\right)\right)=0 \qquad (12)$$

Turunan kedua:

$$\frac{\partial^2 L\left(y,\beta\right)}{\partial^2\beta_k}=-\sum_{i=1}^{n}\frac{x_{ik}^{2}y_i}{\left(\sum_{j=1}^{m}x_{ij}\beta_j\right)^2} \qquad (13)$$

$$\frac{\partial^2 L\left(y,\beta\right)}{\partial\beta_l\partial\beta_k}=-\sum_{i=1}^{n}\frac{x_{ik}x_{il}y_i}{\left(\sum_{j=1}^{m}x_{ij}\beta_j\right)^2} \qquad (14)$$

$$\frac{\partial L\left(y,\beta\right)}{\partial^2\beta_0}=-\sum_{i=1}^{n}\frac{y_i}{(\beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\beta_3x_{3i})^2} \qquad (15)$$

➢ **Estimasi Parameter dengan Pendekatan Model Log Linear**

Model regresi Poisson yang akan diperoleh dengan menggunakan pendekatan model log linear adalah sebagai berikut:

$$\hat{\lambda}=e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}} \qquad (16)$$

Pendugaan parameter dengan menggunakan pendekatan model log linear dengan menggunakan metode MLE untuk model regresi Poisson merupakan turunan pertama dari fungsi log likelihoodnya, maka bentuk yang akan digunakan untuk memperoleh parameter dengan menggunakan metode MLE adalah:

$$L\left(y,\beta\right)=\sum_{i=1}^{n} y_i \ln\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)-\sum_{i=1}^{n}\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)-\sum_{i=1}^{n}\ln\left(y_i!\right) \qquad (17)$$

dan dihasilkan:

$$\sum_{i=1}^{n}\left(\left(\frac{y_i}{\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)}-\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)\right)\left(\frac{\partial\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)}{\partial\hat{\beta}}\right)\right)=0 \qquad (18)$$

Turunan kedua:

$$\frac{\partial^2 L\left(y,\beta\right)}{\partial{\beta_k}^2}=-\sum_{i=1}^{n}x_{ik}^{2}\left(\frac{y_i}{\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)^2}+\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)\right) \qquad (19)$$

$$\frac{\partial^2 L(y,\beta)}{\partial\beta_l\partial\beta_k} = -\sum_{i=1}^{n} x_{il}\, x_{ik}\left(\frac{y_i}{\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)^2} + \left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)\right) \qquad (l \neq k) \tag{20}$$

$$\frac{\partial L(y,\beta)}{\partial\beta_0\partial\beta_k} = -\sum_{i=1}^{n} x_{ik}\left(\frac{y_i}{\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)^2} + \left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)\right) \tag{21}$$

$$\frac{\partial L(y,\beta)}{\partial{\beta_0}^2} = -\sum_{i=1}^{n}\left(\frac{y_i}{\left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)^2} - \left(e^{\hat{\beta}_0+\sum_{j=1}^{m}\hat{\beta}_j x_{ij}}\right)\right) \tag{22}$$

Pada persamaan (12) dan (18) akan digunakan untuk memperoleh $\hat{\beta}$. Perlu dicatat, tidak seperti pada analisis regresi yang biasa, bahwa pada turunan pertama fungsi log-likelihoodnya tidak dapat langsung diperoleh nilai parameternya, sehingga salah satu metode untuk menghitung MLE yaitu dengan menggunakan teknik iteratif.

Beberapa metode yang dapat diusulkan untuk menghitung MLE melalui teknik iteratif ini, diantaranya adalah metode linierisasi, turunan tertejal (*steepest descent*), dan Draper dan Smith (Hajarisman, 2001) mengusulkan metode jalan tengah Marquardt. Selain itu, Myers (1990) mengusulkan untuk menggunakan metode yang disebut metode kuadrat terkecil terbobot ulang secara iteratif (*Iterative reweighted least square,* IRLS) dan metode Gauss-Newton. Sedangkan metode yang lebih umum digunakan pada masalah ini adalah metode Newton-Raphson (Agresti, 1990).

Metode Newton Raphson

Metode Newton Raphson merupakan salah satu metode untuk menghitung MLE melalui teknik iteratif dengan metode pendekatan yang menggunakan satu titik awal.

$$\beta^{(t+1)} = \beta^{(t)} - \left[H^{(t)}\right]^{-1} * g^{(t)}, \tag{23}$$

dimana: $g^{(t)} = \dfrac{\partial L(y,\beta)}{\partial\beta}$

$$H^{(t)} = \frac{\partial^2 L(y,\beta)}{\partial\beta^T\partial\beta}$$

maka bentuk matriks untuk $g^{(t)}$ dan $H^{(t)}$ adalah:

$$g^{(t)} = \begin{bmatrix} \dfrac{\partial L(y,\beta)}{\partial\beta_0} & \dfrac{\partial L(y,\beta)}{\partial\beta_1} & \cdots & \dfrac{\partial L(y,\beta)}{\partial\beta_k} \end{bmatrix} \tag{24}$$

$$H^{(t)} = \begin{bmatrix} \dfrac{\partial^2 L(y,\beta)}{\partial{\beta_0}^2} & \dfrac{\partial^2 L(y,\beta)}{\partial\beta_0\partial\beta_1} & \cdots & \dfrac{\partial^2 L(y,\beta)}{\partial\beta_0\partial\beta_k} \\ \dfrac{\partial^2 L(y,\beta)}{\partial\beta_1\partial\beta_0} & \dfrac{\partial^2 L(y,\beta)}{\partial{\beta_1}^2} & \cdots & \dfrac{\partial^2 L(y,\beta)}{\partial\beta_1\partial\beta_k} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 L(y,\beta)}{\partial\beta_k\partial\beta_0} & \dfrac{\partial^2 L(y,\beta)}{\partial\beta_k\partial\beta_1} & \cdots & \dfrac{\partial^2 L(y,\beta)}{\partial{\beta_k}^2} \end{bmatrix} \tag{25}$$

HASIL DAN PEMBAHASAN

Seperti telah dijelaskan pada pendahuluan, bahwa dalam bahasan ini data yang digunakan sebagai ilustrasi pembentukan model untuk data dengan variabel dependen merupakan data diskrit yang mengukuti distribusi Poisson. Membangkitkan data dengan jumlah sampel 1000, parameter λ dipengaruhi oleh variabel independen X_i

λ_i = exp(b0 +b1X_i) dengan i=1,....1000

Dalam simulasi ini diambil b0=2 dan b1=1. Berdasarkan data (X_i,Yi), dilakukan estimasi parameter menggunakan metode MLE menggunakan pendekatan distribusi normal (regresi linear) dan distribusi Poisson (regresi Poisson). Pada regresi linear data dibangkitkan sedemikian hingga

Var(Y_i) = λ_i

yaitu sama dengan nilai variansi pada distribusi Poisson.

Estimasi Parameter Regresi Poisson

Berdasarkan hasil pada data yang berdistribusi Poisson dengan melakukan replikasi sebanyak 40, maka akan diperoleh nilai variabel independen dan variabel dependen yang akan digunakan untuk mengestimasi parameter dalam regresi poisson. Tabel 1 merupakan tabel hasil estimasi parameter dengan melakukan replikasi sebanyak 40 kali.

Tabel 1. **Estimasi parameter regresi Poisson pada model log linear dengan replikasi sebanyak 40 kali.**

No	b0	b1	No	b0	b1
[1,]	2.014618	0.9899002	[21,]	2.000162	0.9882924
[2,]	1.995027	1.0053512	[22,]	2.011655	0.9880046
[3,]	1.991144	1.0044509	[23,]	2.007278	0.9897750
[4,]	1.997109	1.0071228	[24,]	1.999544	1.0038190
[5,]	1.985287	0.9973325	[25,]	1.994627	1.0003990
[6,]	2.018861	0.9990587	[26,]	2.003672	0.9887621
[7,]	1.996434	1.0099973	[27,]	2.000164	0.9871371
[8,]	2.011502	0.9971834	[28,]	2.003561	0.9862567
[9,]	2.002737	1.0111698	[29,]	1.989819	1.0067965
[10,]	1.986111	1.0049361	[30,]	1.999660	1.0016063
[11,]	2.011638	0.9838045	[31,]	2.011479	1.0062742
[12,]	2.005833	0.9976357	[32,]	2.014921	0.9874263
[13,]	1.995263	1.0091409	[33,]	2.019279	0.9930196
[14,]	1.976698	1.0196407	[34,]	1.992837	1.0081189
[15,]	2.018463	0.9892067	[35,]	2.000685	0.9953407
[16,]	1.969980	1.0170337	[36,]	2.005070	0.9914727
[17,]	2.012856	0.9939624	[37,]	1.996472	1.0036008
[18,]	1.989330	0.9998286	[38,]	1.975403	1.0163725
[19,]	2.003495	1.0061389	[39,]	2.020417	0.9868394
[20,]	1.971552	1.0332100	[40,]	1.997392	0.9971611

Tabel 2. **Estimasi parameter regresi Poisson pada model linear dengan replikasi sebanyak 40 kali.**

No	b0	b1	No	b0	b1
[1,]	6.604480	1.542329	[21,]	6.260227	1.443757
[2,]	4.528289	1.557744	[22,]	4.916360	1.626259
[3,]	5.200574	1.617653	[23,]	6.687097	1.612230
[4,]	5.731523	1.496386	[24,]	4.365954	1.450156
[5,]	4.506369	1.532408	[25,]	6.567917	1.594949
[6,]	3.941255	1.525411	[26,]	4.736484	1.561459
[7,]	5.191849	1.410410	[27,]	5.583958	1.535121
[8,]	4.656280	1.581228	[28,]	3.935584	1.544468

[9,]	3.832811	1.482593	[29,]	4.358800	1.516595
[10,]	5.083398	1.516792	[30,]	4.194426	1.487027
[11,]	4.945988	1.556378	[31,]	3.626556	1.494221
[12,]	5.263396	1.522475	[32,]	4.270232	1.593365
[13,]	4.598323	1.565690	[33,]	4.893617	1.579609
[14,]	6.388321	1.515502	[34,]	5.931982	1.447974
[15,]	5.032204	1.487345	[35,]	5.045957	1.572890
[16,]	5.827288	1.549162	[36,]	4.238662	1.611971
[17,]	3.416215	1.546992	[37,]	4.226404	1.473047
[18,]	7.586075	1.641374	[38,]	3.956635	1.540457
[19,]	4.134141	1.419939	[39,]	4.773847	1.506373
[20,]	5.654318	1.501120	[40,]	4.705784	1.567936

Berdasarkan data simulasi, maka untuk mengetahui nilai estimasi parameter yang mendekati nilai parameter pada model regresi Poisson dengan pendekatan model log linear dan model linear, dapat diuji dengan menggunakan uji *t* dan *Z.* Berikut adalah uji hipotesis yang digunakan untuk mengestimasi parameter dalam regresi Poisson:

Model Log Linear

Pengujian pada β_0

- Hipotesis
  H_0 : $\beta_0 = 2$ melawan H_1 : $\beta_0 \neq 2$
- Menentukan tingkat signifikansi ($\alpha = 0.05$)
- Statistik Uji: Nilai t = –0.024, dengan nilai p-value = 0.981
- Daerah Kritis: H_0 ditolak apabila $t > t_{tabel}$ atau p-value < α
- Keputusan: Tidak tolak H_0 dan dapat disimpulkan bahwa nilai $\beta_0 = 2$.

Pengujian pada β_1

- Hipotesis
  H_0 : $\beta_1 = 1$ melawan H_1 : $\beta_1 \neq 1$
- Menentukan tingkat signifikansi ($\alpha = 0.05$)
- Statistik Uji: Nilai t = 0.0376, dengan nilai p-value = 0.9702
- Daerah Kritis: H_0 ditolak apabila $t > t_{tabel}$ atau p-value < α
- Keputusan: Tidak tolak H_0 dan dapat disimpulkan bahwa nilai β_1=1.

Model Linear

Pengujian pada β_0

- Hipotesis
  H_0 : $\beta_0 = 2$ melawan H_1 : $\beta_0 \neq 2$
- Menentukan tingkat signifikansi ($\alpha = 0.05$)
- Statistik Uji: Nilai t = 59.6966, dengan nilai p-value = 2.2e-16
- Daerah Kritis: H_0 ditolak apabila $t > t_{tabel}$ atau p-value < α
- Keputusan: Tolak H_0 dan dapat disimpulkan bahwa nilai $\beta_0 \neq 2$.

Pengujian pada β_1

- Hipotesis
  H_0 : $\beta_1 = 1$ melawan H_1 : $\beta_1 \neq 1$
- Menentukan tingkat signifikansi ($\alpha = 0.05$)

- Statistik Uji: Nilai t = 19.9047, dengan nilai p-value = 2.2e-16
- Daerah Kritis: H_0 ditolak apabila $t > t_{tabel}$ atau p-value < α
- Keputusan: Tolak H_0 dan dapat disimpulkan bahwa nilai $\beta_1 \neq 1$.

Berdasarkan pengujian parameter dalam model regresi Poisson dengan menggunakan model log linear dan model linear dapat dihasilkan bahwa model linear lebih tepat digunakan dalam mengestimasi parameter regresi Poisson

KESIMPULAN

Dari hasil dan pembahasan dapat diambil kesimpulan sebagai berikut.

1. Penggunaan analisis regresi Poisson lebih baik digunakan dalam pemodelan suatu kasus dimana variabel dependen berbentuk diskrit dan berdistribusi poisson, pada fenomena menyangkut banyaknya suatu kejadian dalam suatu unit dengan batas atas banyaknya kejadian itu tidak terbatas dan merupakan proses kejadian Poisson atau menyangkut dalam distribusi Poisson.
2. Dalam membandingkan model linear dan log linear dengan menggunakan data simulasi diperoleh bahwa model log linear lebih sesuai digunakan pada kasus yang menyangkut proses kejadian Poisson, karena penggunaan pada model linier akan dapat digunakan apabila nilai lambda besar dengan parameter rata – rata diasumsikan bernilai positif.

DAFTAR PUSTAKA

Agresti. 1990. *Categorical Data Analysis*. Second edition. Jhon Willey and sons, Inc, New York.

Draper, N.R., and H. Smith. 1981. *Applied Regression Analysis*. Second edition. John Wiley and Sons. Inc, New York.

Hajarisman, N. 2001. Pengembangan Model Regresi Pada Peubah Respon Diskrit. *Jurnal Forum Teori dan Aplikasi Statistika*, Vol. 1, (Nomor. 1): 15 – 31.

McCullough P. dan Nelder J.A. (1989). *Generalized Linear Models*. 2^{nd}. ed. Chapman and Hall, New York, NY.

Myers. R. H. 1990. *Classical and Modern Regression With Applications*. Penerbit PWS – KENT Publising Company. Boston.

Myers R.H. dan Montgomery D.C., (1997) " A Tutorial On Generalized Linear Model" *Journal of Quality Technologi* Vol. 29. No. 3. pp 274-291.

Nelder, J. A., and R. W. M. Wedderbun. 1972. Generalized Linear Models. *Journal of Royal Statistical Society*, Series A 153: 370–384.

Ratih, R., 2000. Inferensi pada Model – Model Linier Tergeneralisasi, *Majalah Matematika dan Statistika*, Vol. 1. FMIPA Universitas Jember.

Yasin, M., 2003. Pemodelan dan Analisa Data Berkategorik yang Tidak Saling Bebas, *Jurnal ILMU DASAR*, Vol. 4, (Nomor. 2): 11 – 19.