BAB III

MENGATASI OVERDISPERSI PADA REGRESI POISSON DENGAN REGRESI *GENERALIZED* POISSON

3. 1 Regresi Poisson

Regresi Poisson adalah regresi yang dapat digunakan untuk data yang variabel responnya berdistribusi tidak normal dan berjenis diskrit. Syarat utama pada model regresi Poisson yaitu data variabel respon berdistribusi Poisson. Model regresi Poisson dapat dituliskan sebagai berikut:

$$Y_i = E(Y_i) + \varepsilon_i, \quad i = 1,2,\cdots,n \tag{3.1.1}$$

Pada model regresi Poisson terdapat dua *link function* yang biasa digunakan, yaitu *identity link* $g(\mu_i) = \mu_i = \boldsymbol{x'_i\beta}$ dan *ln link* $g(\mu_i) = \ln \mu_i = \boldsymbol{x'_i\beta}$. Jika *identity link* yang digunakan, maka model regresi Poisson dituliskan sebagai berikut:

$$Y_i = \mu_i + \varepsilon_i = \boldsymbol{x'_i\beta} + \varepsilon_i \tag{3.1.2}$$

karena $\mu_i = g^{-1}(\boldsymbol{x'_i\beta}) = \boldsymbol{x'_i\beta}$. Sedangkan jika *ln link* yang digunakan, maka model regresi Poisson dituliskan sebagai berikut:

$$Y_i = \mu_i + \varepsilon_i = \exp(\boldsymbol{x'_i\beta}) + \varepsilon_i \tag{3.1.3}$$

karena $\mu_i = g^{-1}(\ln \boldsymbol{x'_i\beta}) = \exp(\boldsymbol{x'_i\beta})$. *Ln link* lebih baik digunakan untuk model regresi Poisson, karena rata-rata dari variabel responnya akan berbentuk fungsi eksponensial dan menjamin nilainya bernilai positif. Oleh karena itu, regresi Poisson sering juga disebut model ln-linear.

Parameter $\boldsymbol{\beta}$ dalam model regresi Poisson dapat ditaksir dengan menggunakan metode kemungkinan maksimum. Karena *ln link* yang digunakan, maka fungsi peluang dari variabel respon menjadi:

$$f(y;\boldsymbol{\beta}) = \frac{e^{-\exp(\boldsymbol{x'}_i\boldsymbol{\beta})}\exp(\boldsymbol{x'}_i\boldsymbol{\beta})^y}{y!}$$

$$= \frac{e^{-\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_j\right)}\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_j\right)^y}{y!} \qquad (3.1.4)$$

Langkah-langkah penaksiran parameter menggunakan metode kemungkinan maksimum adalah:

- Membentuk fungsi likelihood.

  Fungsi likelihood untuk model regresi Poisson adalah:

$$L(\boldsymbol{\beta};y) = \prod_{i=1}^{n} f\left(y_i | X_{1i}, X_{2i}, \cdots, X_{pi}, \boldsymbol{\beta}\right)$$

$$L(\boldsymbol{\beta};y) = \prod_{i=1}^{n}\left\{\frac{e^{-\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)^{y_i}}{y_i!}\right\} \qquad (3.1.5)$$

- Membentuk fungsi log dari fungsi likelihood yang telah diperoleh. Fungsi log likelihood yang terbentuk dari fungsi likelihood diatas adalah:

$$\ln L(\boldsymbol{\beta}) = \ln\prod_{i=1}^{n}\left\{\frac{e^{-\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)^{y_i}}{y_i!}\right\}$$

$$= \sum_{i=1}^{n}\ln\left\{\frac{e^{-\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)^{y_i}}{y_i!}\right\}$$

$$= \sum_{i=1}^{n}\ln\left[\exp\left(-\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)\right]$$

$$+\sum_{i=1}^{n} y_i \ln\left(\exp\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right) - \sum_{i=1}^{n} \ln(y_i!)$$

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left\{\left(-\exp\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right) + y_i\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right.$$

$$\left. - \, ln(y_i!)\right\} \qquad (3.1.6)$$

- Persamaan log-likelihood diatas didiferensialkan terhadap masing-masing persamaan parameter $\beta_0, \beta_1, \ldots, \beta_p$ dan disamadenganakan nol, sehingga akan diperoleh taksiran parameter yang maksimum atas penyelesaian persamaan tersebut.

$$\mathrm{U}(\boldsymbol{\beta}) = \begin{bmatrix} \dfrac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_0} \\ \dfrac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_1} \\ \vdots \\ \dfrac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_p} \end{bmatrix}$$

$$= \begin{bmatrix} \displaystyle\sum_{i=1}^{n} \left\{ y_i - \exp\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right\} \\ \displaystyle\sum_{i=1}^{n} \left\{ x_{1i}\left(y_i - \exp\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right)\right\} \\ \vdots \\ \displaystyle\sum_{i=1}^{n} \left\{ x_{pi}\left(y_i - \exp\left(\beta_{\mathbf{0}} + \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \beta_{\boldsymbol{j}} \boldsymbol{x}_{\boldsymbol{ji}}\right)\right)\right\} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \mathbf{0}$$

Penentuan penaksir titik dari $\boldsymbol{\beta}$ secara manual mengalami kesulitan, sehingga penentuannya dilakukan dengan menggunakan bantuan *software* statistika, diantaranya SAS, S-PLUS dan Stata. Setelah diperoleh hasil penaksiran parameter, maka taksiran model regresi Poissonnya adalah sebagai berikut:

$$\hat{y}_i = g^{-1}(\boldsymbol{x'}_i\beta) = \exp(\boldsymbol{x'}_i\beta) \qquad (3.1.7)$$

3. 2 Overdispersi

Pada model Poisson berlaku kondisi equidispersi, yaitu rata-rata dan varians dari variabel respon sama. Namun dalam kenyataannya seringkali kondisi tersebut tidak terpenuhi, tidak jarang varians melebihi rata-ratanya. Kondisi tersebut dinamakan overdispersi. Overdispersi dapat diketahui dari hasil pembagian nilai devians dengan derajat bebas. Jika nilainya lebih besar dari satu maka data mengindikasikan gejala overdispersi.

Overdispersi dapat terjadi karena adanya sumber keragaman yang tidak teramati, pada variabel respon ada data yang hilang, adanya pencilan pada data, adanya interaksi antar pengamatan dan kesalahan spesifikasi *link function*. Jika model Poisson yang mengindikasikan gejala overdispersi tetap digunakan untuk menggambarkan model data, maka taksiran parameter regresinya tetap konsisten tetapi tidak efisien. Akibatnya, taksiran parameter menjadi bias sehingga mengakibatkan kesalahan penarikan kesimpulan mengenai data yang diamati.

Oleh karena itu diperlukan sebuah model lain yang dapat mengatasi gejala overdispersi, yaitu model *generalized* Poisson.

3. 3 Regresi *Generalized* Poisson

Distribusi *generalized* Poisson adalah perluasan dari distribusi Poisson. Sebuah sebaran data dikatakan berdistribusi *generalized* Poisson jika fungsi peluangnya berbentuk:

$$f(y) = \left(\frac{\mu_i}{1+k\mu_i}\right)^{y_i}\frac{(1+ky_i)^{y_i-1}}{y_i!}\exp\left(-\frac{\mu_i(1+ky_i)}{(1+k\mu_i)}\right);\ y_i = 0,1,2,\dots \quad (3.3.1)$$

dengan rata-rata dan variansnya adalah μ_i dan $\mu_i(1+k\mu_i)^2$. Untuk mencari rata-rata dan varians dari distribusi *generalized* Poisson, akan dibuktikan terlebih dahulu bahwa distribusi *generalized* Poisson termasuk dalam distribusi keluarga eksponensial. Fungsi peluang dari *generalized* Poisson dapat dituliskan dalam bentuk:

$$f(y) = \exp\left\{y_i\ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{\mu_i(1+ky_i)}{(1+k\mu_i)} + (y_i-1)\ln(1+ky_i) - \ln(y_i!)\right\}$$

$$= \exp\left\{y_i\ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{\mu_i}{(1+k\mu_i)} - \frac{k\mu_i y_i}{(1+k\mu_i)} + (y_i-1)\ln(1+ky_i) - \ln(y_i!)\right\}$$

$$f(y) = \exp\left\{y_i\left(\ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{k\mu_i}{(1+k\mu_i)}\right) - \frac{\mu_i}{(1+k\mu_i)} + (y_i-1)\ln(1+ky_i) - \ln(y_i!)\right\} \quad (3.3.2)$$

Dari penulisan bentuk lain di atas dapat dilihat bahwa *generalized* Poisson termasuk dalam distribusi keluarga eksponensial dengan

$$\theta = \ln\left(\frac{\mu_i}{1 + k\mu_i}\right) - \frac{k\mu_i}{(1 + k\mu_i)}$$

$$b(\theta) = \frac{\mu_i}{(1 + k\mu_i)}$$

$$\alpha(\phi) = 1$$

$$h(y, \phi) = (y_i - 1)\ln(1 + ky_i) - \ln(y_i!)$$

Telah dibuktikan pada Lampiran 2.1 dan Lampiran 2.2 bahwa rata-rata dan varians dari keluarga eksponensial adalah $b'(\theta)$ dan $b"(\theta)\alpha(\phi)$. Sehingga akan dicari $b'(\theta)$ dan $b"(\theta)\alpha(\phi)$ untuk distribusi *generalized* Poisson.

Diketahui

$$\theta = \ln\left(\frac{\mu_i}{1 + k\mu_i}\right) - \frac{k\mu_i}{(1 + k\mu_i)}$$

dan

$$b(\theta) = \frac{\mu_i}{(1 + k\mu_i)}$$

sehingga

$$\theta = \ln\big(b(\theta)\big) - kb(\theta)$$

$$\frac{\partial\theta}{\partial\theta} = \frac{1}{b(\theta)} b^{'}(\theta) - kb'(\theta)$$

$$1 = \frac{b^{'}(\theta) - kb'(\theta)b(\theta)}{b(\theta)}$$

$$b(\theta) = b^{'}(\theta) - kb'(\theta)b(\theta)$$

$$b(\theta) = b^{'}(\theta)[1 - kb(\theta)]$$

$$b^{'}(\theta) = \frac{b(\theta)}{1 - kb(\theta)}$$

$$= \frac{\frac{\mu_i}{(1+k\mu_i)}}{1-k\frac{\mu_i}{(1+k\mu_i)}}$$

$$= \frac{\frac{\mu_i}{(1+k\mu_i)}}{\frac{(1+k\mu_i)-k\mu_i}{(1+k\mu_i)}}$$

$$b^{'}(\theta) = \mu_i$$

Karena distribusi *generalized* Poisson termasuk dalam distribusi keluarga eksponensial dan $b^{'}(\theta) = \mu_i$, maka terbukti bahwa rata-rata dari distribusi *generalized* Poisson adalah μ_i. Dari proses pembuktian rata-rata di atas diperoleh

$$1 = \frac{b^{'}(\theta) - kb'(\theta)b(\theta)}{b(\theta)}$$

$$= \frac{b^{'}(\theta)}{b(\theta)} - kb'(\theta)$$

jika didiferensialkan terhadap θ akan menjadi

$$0 = \frac{b^{''}(\theta)b(\theta) - \left(b^{'}(\theta)\right)^2}{\left(b(\theta)\right)^2} - kb''(\theta)$$

$$0 = \frac{b^{''}(\theta)b(\theta) - \left(b^{'}(\theta)\right)^2 - kb''(\theta)\left(b(\theta)\right)^2}{\left(b(\theta)\right)^2}$$

$$b^{''}(\theta)b(\theta) - \left(b^{'}(\theta)\right)^2 - kb^{''}(\theta)\left(b(\theta)\right)^2 = 0$$

$$b^{''}(\theta)b(\theta) - kb''(\theta)\left(b(\theta)\right)^2 = \left(b^{'}(\theta)\right)^2$$

$$b^{''}(\theta)\left[b(\theta) - k\left(b(\theta)\right)^2\right] = \left(b^{'}(\theta)\right)^2$$

$$b^{''}(\theta) = \frac{\left(b^{'}(\theta)\right)^2}{b(\theta) - k\left(b(\theta)\right)^2}$$

$$b^{''}(\theta) = \frac{(\mu_i)^2}{\frac{\mu_i}{(1+k\mu_i)} - k\left(\frac{\mu_i}{(1+k\mu_i)}\right)^2}$$

$$b^{''}(\theta) = \frac{(\mu_i)^2}{\frac{\mu_i}{(1+k\mu_i)} - k\frac{\mu_i{}^2}{(1+k\mu_i)^2}}$$

$$b^{''}(\theta) = \frac{(\mu_i)^2}{\frac{\mu_i(1+k\mu_i) - k\mu_i{}^2}{(1+k\mu_i)^2}}$$

$$= \frac{(\mu_i)^2}{\frac{\mu_i + k\mu_i{}^2 - k\mu_i{}^2}{(1+\alpha\mu_i)^2}}$$

$$= (\mu_i)^2 \frac{(1+k\mu_i)^2}{\mu_i}$$

$$b^{''}(\theta) = \mu_i(1+k\mu_i)^2$$

Karena $b^{''}(\theta)a(\phi) = \mu_i(1+k\mu_i)^2.1 = \mu_i(1+k\mu_i)^2$, maka terbukti bahwa varians dari distribusi *generalized* Poisson adalah $\mu_i(1+k\mu_i)^2$.

Fungsi peluang distribusi *generalized* Poisson mempunyai dua parameter yaitu μ sebagai parameter natural dan k sebagai parameter dispersi. Ketika parameter dispersi k bernilai nol, maka fungsi peluang distribusi *generalized* Poisson akan menjadi fungsi peluang distribusi Poisson. Jika parameter dispersi $k < 0$, maka varians akan lebih kecil dari rata-rata dan disebut dengan kondisi underdispersi, sedangkan jika parameter dispersi $k > 0$, maka varians akan melebihi rata-rata yang disebut dengan kondisi overdispersi.

Karena $\mu_i = \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)$, maka fungsi peluang distribusi *generalized* Poisson menjadi:

$$f(y) = \left(\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}\right)^{y_i} \frac{(1 + ky_i)^{y_i - 1}}{y_i!}$$

$$\exp\left(-\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)(1 + ky_i)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)}\right);$$

$$y_i = 0,1,2,\ldots \qquad (3.3.3)$$

Parameter $\boldsymbol{\beta}$ dan k dapat ditaksir dengan metode kemungkinan maksimum. Langkah penaksiran parameter dalam regresi *generalized* Poisson sama dengan lagkah penaksiran parameter dalam model regresi Poisson, yaitu:

- Membentuk fungsi likelihood. Fungsi likelihood untuk model regresi *generalized* Poisson adalah:

$$L(\boldsymbol{\beta}, k; y) = \prod_{i=1}^{n} f(y; \boldsymbol{\beta}, k)$$

$$= \prod_{i=1}^{n} \left\{ \left(\frac{\mu_i}{1 + k\mu_i}\right)^{y_i} \frac{(1 + ky_i)^{y_i - 1}}{y_i!} \exp\left(-\frac{\mu_i(1 + ky_i)}{(1 + k\mu_i)}\right) \right\}$$

$$L(\boldsymbol{\beta}, k; y) = \prod_{i=1}^{n} \left\{ \left(\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}\right)^{y_i} \frac{(1 + ky_i)^{y_i - 1}}{y_i!} \right.$$

$$\left. \exp\left(-\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)(1 + ky_i)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)}\right) \right\} \qquad (3.3.4)$$

- Membentuk fungsi log dari fungsi likelihood yang telah diperoleh. Fungsi log likelihood yang terbentuk dari fungsi likelihood diatas adalah:

$$\ln L(\boldsymbol{\beta},\mathrm{k}) = \ln \prod_{i=1}^{n}\left\{\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right)^{y_i}\frac{(1+ky_i)^{y_i-1}}{y_i!}\right.$$

$$\left.\exp\left(-\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)(1+ky_i)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)}\right)\right\}$$

$$= \sum_{i=1}^{n}\ln\left\{\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right)^{y_i}\frac{(1+ky_i)^{y_i-1}}{y_i!}\right.$$

$$\left.+\exp\left(-\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)(1+ky_i)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)}\right)\right\}$$

$$= \sum_{i=1}^{n}\left\{\ln\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right)^{y_i}+\ln\left(\frac{(1+ky_i)^{y_i-1}}{y_i!}\right)\right.$$

$$\left.+\ln\left(\exp\left(-\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)(1+ky_i)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)}\right)\right)\right\}$$

$$= \sum_{i=1}^{n}\left\{y_i\ln\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right)+(y_i-1)\ln(1+ky_i)\right.$$

$$\left.-\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)(1+ky_i)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)}-\ln(y_i!)\right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i \ln \left(\exp \left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji} \right) \right) \right.$$

$$- y_i \ln \left(1 + k \exp \left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji} \right) \right)$$

$$+ (y_i - 1) \ln(1 + k y_i)$$

$$\left. - \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)(1 + k y_i)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)} - \ln(y_i!) \right\}$$

$$\ln L(\boldsymbol{\beta}, \mathrm{k}) = \sum_{i=1}^{n} \left\{ y_i \left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji} \right) - y_i \ln \left(1 + k \exp \left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji} \right) \right) \right.$$

$$+ (y_i - 1) \ln(1 + k y_i) - \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)(1 + k y_i)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)}$$

$$\left. - \ln(y_i!) \right\} \tag{3.3.5}$$

- Terdapat tiga langkah dalam menaksir parameter regresi *generalized* Poisson, yaitu mendiferensialkan persamaan log-likelihood diatas terhadap β_0 kemudian disamakan dengan nol, mendiferensialkan persamaan log-likelihood diatas terhadap β_j ; $j = 1,2, \dots$ kemudian disamakan dengan nol, dan mendiferensialkan persamaan log-likelihood terhadap k kemudian di sama dengankan nol. Maka akan diperoleh persamaan:

- $\frac{\partial \ln L(\boldsymbol{\beta},k)}{\partial \beta_0} =$

$$\sum_{i=1}^{n} y_i - y_i \frac{k \exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)} -$$

$$\frac{(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)-k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2}$$

$$= \sum_{i=1}^{n} y_i - y_i \frac{k \exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}$$

$$- \frac{(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\left[1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)-k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right]}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2}$$

$$= \sum_{i=1}^{n} \left\{ y_i \left(1 - \frac{k \exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right) - \frac{(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i \left(\frac{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right) - k \exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}\right) - \frac{(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{y_i}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)} - \frac{(1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left(\frac{y_i\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right) - (1+ky_i)\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{y_i + ky_i \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}) - \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}) - ky_i \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{(y_i - \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}))}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{(1 + k\mu_i)^2} \qquad (3.3.6)$$

$$= 0$$

- $\frac{\partial \ln L(\boldsymbol{\beta},k)}{\partial \beta_j} =$

$$\sum_{i=1}^{n} \left\{ y_i x_{ji} - y_i x_{ji} \frac{k \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{1 + k(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})} - \right.$$

$$\left. \frac{(1+ky_i)x_{ji} \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})(1 + k\exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})) - k\, x_{ji} \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})(1+ky_i) \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{\left(1 + k\exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ji} - y_i x_{ji} \frac{k \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{1 + k\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} \right.$$

$$\left. - \frac{(1 + ky_i) x_{ji} \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}) \left[1 + k \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}) - k \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})\right]}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ji} \left(1 - \frac{k \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} \right) \right.$$

$$\left. - \frac{(1 + ky_i) x_{ji} \exp(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji})}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ji} \left(\frac{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right) - k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} \right) \right.$$

$$\left. - \frac{(1 + ky_i) x_{ji} \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{y_i x_{ji}}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} - \frac{(1 + ky_i) x_{ji} \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left(\frac{y_i x_{ji}\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right) - (1 + ky_i) x_{ji} \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{y_i x_{ji} + y_i x_{ji} k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right) - x_{ji}\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right) - k y_i x_{ji} \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{x_{ji}\left(y_i - \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \frac{x_{ji}(y_i - \mu_i)}{(1 + k\mu_i)^2}, \quad j = 1,2, \dots \tag{3.3.7}$$

$$= 0$$

- $\frac{\partial \ln L(\beta,k)}{\partial k} =$

$$\sum_{i=1}^{n} \left\{ - \frac{y_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} + \frac{y_i(y_i - 1)}{1 + ky_i} - \right.$$

$$\left. \frac{y_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right) - \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)(1 + ky_i)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} + \frac{y_i(y_i - 1)}{1 + ky_i} \right.$$

$$\left. - \frac{y_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right) + ky_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)^2 - \left(\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)^2 - ky_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)^2\right)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)}{1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)} + \frac{y_i(y_i - 1)}{1 + ky_i} \right.$$

$$\left. - \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\left(y_i - \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ji}\right)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \mu_i}{1 + k\mu_i} + \frac{y_i(y_i - 1)}{1 + ky_i} - \frac{\mu_i(y_i - \mu_i)}{(1 + k\mu_i)^2} \right\} \tag{3.3.8}$$

$$= 0$$

Setelah penaksiran parameter selesai, taksiran model regresi *generalized* Poissonnya adalah:

$$\hat{y}_i = \exp(\boldsymbol{x'}_i \boldsymbol{\beta})$$

3. 4 Pengujian Dalam Model Regresi

Hasil penaksiran parameter regresi yang diperoleh harus diuji terlebih dahulu apakah parameter tersebut memiliki keberartian terhadap variabel respon ataupun tidak. Maka akan dilakukan terlebih dahulu uji keberartian koefisein dengan langkah pengujiannya sebagai berikut:

- Hipotesis pengujian:

  $H_0 : \beta_j = 0$

  $H_1 : \beta_j \neq 0$

- Statistik uji:

$$w = \left(\frac{\beta_j}{s_e(\beta_j)} \right)^2$$

- Kriteria: tolak H_0 jika $w > \chi^2{}_{(\alpha,1)}$ atau tolak H_0 jika nilai $p - value > \alpha$.

Setelah diperoleh model regresi Poisson dan regresi *Generalized* Poisson dari hasil uji keberartian koefisein di atas, akan dilihat kecocokan model regresinya. Terdapat beberapa metode yang bisa digunakan untuk membandingkan kecocokan kedua model tersebut dalam menggambarkan data.

3.4.1 Uji Rasio Likelihood

Untuk menguji model regresi *generalized* Poisson lebih baik digunakan dibanding model regresi Poisson yaitu dengan menguji nilai parameter dispersi k. Hipotesis pengujiannya sebagai berikut:

$H_0 : k = 0$

$H_1 : k \neq 0$

dengan statistik uji

$$T = -2(\ell_1 - \ell_0)$$

dimana ℓ_1 adalah nilai log-likelihood pada model regresi *generalized* Poisson dan ℓ_0 adalah nilai log-likelihhod pada regresi model Poisson. Kriteria pengujian uji rasio likelihood ini adalah tolak H_0 jika $T > \chi^2_{(\alpha,1)}$.

Penolakan H_0 menunjukkan bahwa regresi *generalized* Poisson lebih baik digunakan daripada regresi Poisson biasa.

3.4.2 Akaike Information Criterion

Menurut Bozdogan (2000:63), AIC didefinisikan sebagai berikut:

$$AIC = -2\ell + 2p$$

dengan ℓ adalah nilai ln-likelihood dari model dan p adalah banyaknya parameter. Semakin kecil nilai AIC maka model semakin baik.