BAB II
LANDASAN TEORI

2.1 Distribusi Poisson

Menurut Harinaldi (2005: 87), dalam eksperimen Poisson, probabilitas memperoleh dengan tepat peristiwa Y sebanyak $y$kejadian untuk setiap satu satuan unit (waktu dan ruang) yang ditentukan membentuk sebuah distribusi yang fungsi probabilitasnya adalah

$$f(y,\lambda) = \Pr(Y = y;\lambda) = \frac{\lambda^{y} e^{-\lambda}}{y!} \qquad y = 0,1,2,... \qquad (2.1)$$

dimana

λ : laju kejadian (rata-rata banyaknya kejadian dalam satu satuan unit tertentu)

$e = 2{,}71828 \ldots$,

Dalam kehidupan sehari-hari kasus data berdistribusi Poisson merupakan data yang jarang terjadi dalam selang waktu tertentu.

2.2 Regresi poisson

Menurut Safrida, Ispriyanti, dan Widiharih (2013) regresi Poisson merupakan salah satu regresi nonlinier yang sering digunakan untuk memodelkan hubungan antara variabel respon yang berupa data diskrit dengan variabel prediktor yang berupa data diskrit atau kontinu. Regresi Poisson merupakan penerapan dari *Generalized Linear Model (GLM).Generalized Linear Model (GLM)* merupakan perluasan dari model regresi umum untuk variabel respon yang memiliki sebaran eksponensial. Regresi Poisson digunakan untuk menganalisis data count (berjenis diskrit). Pada regresi Poisson diasumsikan variabel respon Yberdistribusi Poisson dan tidak terjadi multikolinearitas diantara masing-masing variabel prediktor (X).

Dalam regresi Poisson terdapat asumsi yang harus dipenuhi yaitu variabel respon (Y) diskrit dan asumsi equidispersi. Equidispersi yaitu nilai rata-

rata sama dengan nilai varian atau $(Y|x) = E(Y|x) = \mu$. Regresi Poisson ada 2 tipe yaitu regresi Poisson sederhana dan regresi Poisson berganda.

Teorema

Misal Yvariabel random berdistribusi Poisson dengan parameter λmaka rata-rata dan variansi Y adalah λ.

Bukti

Rata-rata dari Y yang berdistribusi Poisson dengan parameter λadalah

$$E(Y) = \sum_{y=0}^{\infty} y.f(y,\lambda)$$

$$= \sum_{y=0}^{\infty} y.\frac{\lambda^y e^{-\lambda}}{y!}$$

$$= \sum_{y=1}^{\infty} \frac{\lambda^y e^{-\lambda}}{(y-1)!}$$

$$= \sum_{y=1}^{\infty} \frac{\lambda\lambda^{(y-1)} e^{-\lambda}}{(y-1)!}$$

$$= \lambda\sum_{y=1}^{\infty} \frac{\lambda^{(y-1)} e^{-\lambda}}{(y-1)!}$$

$$= \lambda\sum_{z=0}^{\infty} \frac{\lambda^z e^{-\lambda}}{z!} \qquad \text{(misalkan } z = y-1, y=1 \text{ maka } z=0\text{)}$$

$$= \lambda\sum_{z=0}^{\infty} f(z;\lambda)$$

$$= \lambda$$

Sedangkan variansi dari Y yang berdistribusi Poisson dengan parameter λ adalah

$$Var(Y) = E(Y^2) - [(E(Y)]^2$$

$$= E(Y^2 - Y + Y) - [(E(Y)]^2$$

$$= E(Y(Y-1) + Y) - [(E(Y)]^2$$

$$= E\big(Y(Y-1)\big) + E(Y) - [(E(Y)]^2 \qquad \text{(Dari sifat ekspektasi)}$$

$$= \sum_{y=0}^{\infty} y(y-1).\frac{\lambda^y e^{-\lambda}}{y!} + E(Y) - [(E(Y)]^2$$

$$= \sum_{y=0}^{\infty} \frac{\lambda^y e^{-\lambda}}{(y-2)!} + E(Y) - [(E(Y)]^2$$

$$= \sum_{y=2}^{\infty} \frac{\lambda^2 \lambda^{(Y-2)} e^{-\lambda}}{(y-2)!} + E(Y) - [(E(Y)]^2$$

$$= \lambda^2 \sum_{y=2}^{\infty} \frac{\lambda^Z e^{-\lambda}}{Z!} + E(Y) - [(E(Y)]^2$$

Misalkan $z = y - 2,\ y = 2$maka $z = 0$

$$Var(Y) = \lambda^2 . 1 + \lambda - \lambda^2 = \lambda$$

2.2.1 Analisis Regresi Poisson Sederhana

Menurut Hertriyanti (2006: 15) analisis regresi Poisson sederhana adalah sebuah metode statistika yang digunakan untuk menganalisis hubungan antara sebuah variabel respon (Y) yang menyatakan data diskrit dan sebuah variabel prediktor (X). Variabel respon (Y) diberikan $X = x$ diasumsikan berdistribusi Poisson, sedangkan variabel prediktor (X) dapat berjenis diskrit, kontinu atau berjenis kategorik.

Misalkan ingin diketahui hubungan antara sebuah variabel respon (Y) yang menyatakan data diskrit dan sebuah variabel prediktor (X) yang berjenis kontinu atau kategorik. Variabel random respon Ydiberikan $X = x$ diasumsikan berdistribusi Poisson. Jika diberikan sebuah sampel berisi n buah pasangan pengamatan yang saling bebas, yaitu $\{(X_i, Y_i)\ ; i = 1,2,3, \dots, n\}$ dengan $X_i$dan Y_i berturut-turut adalah pengamatan ke-i dari variabel$X$ dan Y, maka hubungan antara Ydan X tidak dapat dijelaskan oleh model regresi linear sederhana.

$$P(Y|\mu) = \frac{\mu^{yi} e^{-\mu}}{y!}\ , y = 0,1,2, \dots$$

$$E(Y) = \mu$$

$$\mu = \beta_0 + \beta_1 X_i ; i = 1,2, \dots, n$$

Maka model regresi poisson sederhana yaitu:

$$\ln\big(\mu_i(x_i)\big) = \beta_0 + \beta_1 X_i ; i = 1,2, \dots, n \qquad (2.2)$$

Atau equivalen dengan

$$\mu_i(x_i) = e^{\beta_0 + \beta_1 X_i} ; i = 1,2, \dots, n \qquad (2.3)$$

Dengan β_0, β_1 adalah parameter yang tidak diketahui. Model (2.2) sering disebut dengan fungsi penghubung logaritma atau fungsi log *link*.

2.2.2 Analisis Regresi Poisson Berganda

Menurut Hertriyanti (2006:35) regresi Poisson berganda adalah regresi yang menganalisis hubungan antara sebuah variabel respon (Y) yang merupakan data berjenis diskrit, diasumsikan berdistribusi Poisson dan p buah variabel prediktor $(X) x_1, x_2, \dots, x_p$ yang berjenis diskrit, kontinu atau kategorik.

Model regresi Poisson berganda merupakan perluasan dari model regresi Poisson sederhana, dimana dalam regresi Poisson berganda akan diketahui hubungan antara sebuah variabel respon (Y) yang berjenis diskrit dan p buah variabel prediktor $X_1, X_2, \dots, X_p$yang berjenis kontinu atau kategorik.Variabel random dari respon Y diberikan $X_1 = x_1,\ X_2 = x_2, \dots, X_p = x_p$yang diasumsikan berdistribusi Poisson.

Diberikan sebuah sampel yang berisi n buah pasangan pengamatan yang saling bebas $\{(X_{1i},\ X_{2i}, \dots, X_{pi}, Y_i); i = 1,2, \dots, n\}$ dengan $X_{1i},\ X_{2i}, \dots, X_{pi}$ berurut turut adalah pengamatan ke-i dari variabel$X_1,\ X_2, \dots, X_p$ dan Y_i adalah pengamatan ke-i dari variabel Y. Jika rata-rata bersyarat dari Y_i diberikan nilai $X_{1i},\ X_{2i}, \dots, X_{pi}$ dinyatakan oleh

$$E\big(Y_i \big| X_{1i} = x_{1i}, X_{2i} = x_{2i}, \dots, X_{pi} = x_{pi}\big) = \mu_i\big(x_{1i},\ x_{2i}, \dots, x_{pi}\big),\ \ i = 1,2, \dots, n$$

$$\mu_i\big(x_{1i},\ x_{2i}, \dots, x_{pi}\big) = e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}}; \qquad i = 1,2, \dots, n \quad (2.4)$$

Maka model regresi poisson bergandanya adalah

$$\ln\big(\mu_i(x_i)\big) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} \qquad (2.5)$$

Dengan $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ menyatakan parameter-parameter yang tidak diketahui.

Model (2.3) ekuivalen dengan $\mu_i\big(x_{1i},\ x_{2i}, \dots, x_{pi}\big) = e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}}$

Dengan $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ menyatakan parameter-parameter yang tidak diketahui.

Model (2.3) ekuivalen dengan

$$\begin{aligned} \hat{\mu}_i\big(x_{1i},\ x_{2i}, \dots, x_{pi}\big) &= e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}} \\ &= e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})} \qquad (2.6) \end{aligned}$$

2.2.3 Penaksiran Parameter Pada Model Regresi Poisson

Pada model regresi Poisson harus dilakukan penaksiran pada$\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ dengan $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$adalah parameter yang tidak diketahui. Metode yang digunakan untuk menaksir parameter $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ yaitu metode *Maximum Likelihood Estimation* (MLE). Penaksiran parameter dilakukan pada semua model regresi Poisson, baik sederhana maupun berganda.

Jika diberikan sebuah sampel berisi n buah pasangan pengamatan yang saling bebas, yaitu $\{(X_{1i},\ X_{2i}, \ldots, X_{pi}, Y_i); i = 1,2, \ldots, n\}$dengan$X_i$ danY_i berturut-turut adalah pengam-atan ke-i dari variabel X dan Y dan asumsi untuk setiap $X_{1i} = x_{1i},\ X_{2i} = x_{2i}, \ldots, X_{pi} = x_{pi}$distribusi dari Y_i adalah Poisson dan $E(Y_i|X_i = x_i) = \mu_i(x_i)$, maka fungsi probabilitas bersyarat dari $Y_i$oleh $x_{1i},\ x_{2i}, \ldots, x_{pi}$adalah

$$f\left(y_i \middle| x_{1i}, x_{2i}, \ldots, x_{pi};\ \mu_i(x_{1i}, x_{2i}, \ldots, x_{pi})\right) = \frac{[\mu_i(x_{1i}, x_{2i}, \ldots, x_{pi})]^{y_i} e^{[-\mu_i(x_{1i}, x_{2i}, \ldots, x_{pi})]}}{y_i!}\ ,$$

$$y_i = 0,1,2, \ldots$$

Karena $\mu_i(x_{1i}, x_{2i}, \ldots, x_{pi}) = [e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})}]$ maka diperoleh

$$f(y_i|x_{1i}, x_{2i}, \ldots, x_{pi};\ \boldsymbol{\beta}) = \frac{[e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})}]^{y_i} e^{-[e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})}]}}{y_i!} \qquad (2.7)$$

Dalam hal ini $\boldsymbol{\beta} = \beta_0, \beta_1, \beta_2, \ldots, \beta_p$.

Fungsi *likelihood* dapat diperoleh dengan mengalikan semua fungsi probabilitas bersyarat dari Y_i oleh x_i sehingga

$$K(\boldsymbol{\beta}) = \prod_{i=1}^{n} f(y_i|x_{1i}, x_{2i}, \ldots, x_{pi};\ \boldsymbol{\beta})$$

$$= \prod_{i=1}^{n} \left\{ \frac{[e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})}]^{y_i} e^{-[e^{(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ji})}]}}{y_i!} \right\} \qquad (2.8)$$

Agar persamaan (2.8) mudah diselesaikan maka diubah bentuk menjadi fungsi *loglikelihood*. Fungsi *log likelihood*nya yaitu $k(\boldsymbol{\beta}) = \log K(\boldsymbol{\beta})$ sehingga diperoleh

$$k(\boldsymbol{\beta}) = \log K(\boldsymbol{\beta})$$

$$= \log\left\{\prod_{i=1}^{n}\left\{\frac{[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]^{y_i} e^{-[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]}}{y_i!}\right\}\right\}$$

$$= \sum_{i=1}^{n}\log\prod_{i=1}^{n}\left\{\frac{[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]^{y_i} e^{-[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]}}{y_i!}\right\}$$

$$= \sum_{i=1}^{n}\left\{\log[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]^{y_i} + \log e^{-[e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}]} - \log y_i!\right\}$$

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

Dalam persamaan tersebut x_i dan y_i bilangan yang berasal dari pengamatan sedangkan$(\beta_0 + \sum_{j=1}^{p}\beta_j X_{ji})$ dianggap berubah bila garis regresinya berubah (Sembiring, 1995: 40). Dari segi kalkulus, ini berarti bahwa perlu dicari turunan dari fungsi *log likelihood*terhadap $k(\boldsymbol{\beta})$ kemudian menyamakannya dengan nol sehingga diperoleh nilai p +1 persamaan *likelihood*.

Turunan pertama dari $k(\boldsymbol{\beta})$ terhadap β_0 yaitu

$$R_0(\boldsymbol{\beta}) = \frac{\partial k(\boldsymbol{\beta})}{\partial \beta_0} = \sum_{i=1}^{n}\left\{y_i - e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\} = 0$$

$$R_1(\boldsymbol{\beta}) = \frac{\partial k(\boldsymbol{\beta})}{\partial \beta_1} = \sum_{i=1}^{n}\left\{y_i x_{1i} - x_{1i} e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}\right\} = 0$$

$$= \sum_{i=1}^{n}\left\{x_{1i}(y_i - e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\} = 0$$

$$\dots \tag{2.9}$$

$$R_p(\boldsymbol{\beta}) = \frac{\partial k(\boldsymbol{\beta})}{\partial \beta_p} = \sum_{i=1}^{n}\left\{y_i x_{pi} - x_{pi} e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}\right\} = 0$$

$$= \sum_{i=1}^{n}\left\{x_{pi}(y_i - e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\} = 0$$

Bentuk vektor dari persamaan (2.9) yaitu

$$R(\beta)=\begin{bmatrix}R_0(\boldsymbol{\beta})\\R_1(\boldsymbol{\beta})\\\vdots\\R_p(\boldsymbol{\beta})\end{bmatrix}=\begin{bmatrix}\dfrac{\partial k(\boldsymbol{\beta})}{\partial\beta_0}\\\dfrac{\partial k(\boldsymbol{\beta})}{\partial\beta_1}\\\vdots\\\dfrac{\partial k(\boldsymbol{\beta})}{\partial\beta_p}\end{bmatrix}=\begin{bmatrix}\displaystyle\sum_{i=1}^{n}\left\{y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}\\\displaystyle\sum_{i=1}^{n}\left\{x_{1i}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\\\vdots\\\displaystyle\sum_{i=1}^{n}\left\{x_{pi}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=0$$

Selajutnya nilai dari $\boldsymbol{\beta}=\begin{bmatrix}\beta_0\\\beta_1\\\vdots\\\beta_p\end{bmatrix}$kemudian akan memaksimumkan $k(\boldsymbol{\beta})$.

Nilai taksiran maksimum likelihood dinotasikan dengan $\hat{\beta}=\begin{bmatrix}\beta_0\\\beta_1\\\vdots\\\beta_p\end{bmatrix}$.

Turunan kedua atau matriks hessian dari $k(\boldsymbol{\beta})$ terhadap β_0 yaitu

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

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_1\partial\beta_0}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}\right\}}{\partial\beta_1}=-\sum_{i=1}^{n}\left\{x_{1i}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_0\partial\beta_1}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{x_{1i}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\right\}}{\partial\beta_0}=-\sum_{i=1}^{n}\left\{x_{1i}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_p\partial\beta_0}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{x_{pi}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\right\}}{\partial\beta_p}=-\sum_{i=1}^{n}\left\{x_{pi}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_p\partial\beta_1}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{x_{1i}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\right\}}{\partial\beta_p}=-\sum_{i=1}^{n}\left\{x_{1i}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}x_{pi}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_0\partial\beta_p}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{x_{pi}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\right\}}{\partial\beta_0}=-\sum_{i=1}^{n}\left\{x_{pi}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial\beta_1\partial\beta_p}=\frac{\partial\left\{\sum_{i=1}^{n}\left\{x_{pi}(y_i-e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})})\right\}\right\}}{\partial\beta_1}=-\sum_{i=1}^{n}\left\{x_{pi}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}x_{1i}\right\}$$

$$\frac{\partial^2 k(\boldsymbol{\beta})}{\partial {\beta_p}^2} = \frac{\partial \left\{\sum_{i=1}^{n}\left\{x_{pi}\left(y_i - e^{(\beta_0+\sum_{j=1}^{p}\beta_j X_{ji})}\right)\right\}\right\}}{\partial \beta_p} = -\sum_{i=1}^{n}\left\{x_{pi}e^{(\beta_0+\sum_{j=1}^{p}\beta_j x_{ji})}x_{pi}\right\}$$

Jika dibuat bentuk matriks menjadi

$$H(\boldsymbol{\beta}) = \begin{bmatrix} H_{00}(\boldsymbol{\beta}) & H_{01}(\boldsymbol{\beta}) & \dots & H_{0p}(\boldsymbol{\beta}) \\ H_{10}(\boldsymbol{\beta}) & H_{11}(\boldsymbol{\beta}) & \dots & H_{1p}(\boldsymbol{\beta}) \\ \vdots & \vdots & \vdots & \vdots \\ H_{p0}(\boldsymbol{\beta}) & H_{p1}(\boldsymbol{\beta}) & \dots & H_{pp}(\boldsymbol{\beta}) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \dots & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_0} \\ \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_1^2} & \dots & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_p} & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} & \dots & \frac{\partial^2 k(\boldsymbol{\beta})}{\partial \beta_p^2} \end{bmatrix}$$

$$= \begin{bmatrix} -\sum_{i=1}^{n}\left\{e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & -\sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & \dots & -\sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} \\ -\sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & -\sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{1i}\right\} & \dots & -\sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{pi}\right\} \\ \vdots & \vdots & \vdots & \vdots \\ -\sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & -\sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{1i}\right\} & \dots & -\sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{pi}\right\} \end{bmatrix}$$

$$= -\begin{bmatrix} \sum_{i=1}^{n}\left\{e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & \sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & \dots & \sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} \\ \sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & \sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{1i}\right\} & \dots & \sum_{i=1}^{n}\left\{x_{1i}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{pi}\right\} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}\right\} & \sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{1i}\right\} & \dots & \sum_{i=1}^{n}\left\{x_{pi}e^{\left(\beta_0+\sum_{j=1}^{\infty}\beta_j x_{ji}\right)}x_{pi}\right\} \end{bmatrix}$$

(2.10)

Persamaan (2.10) merupakan persamaan dalam bentuk eksponensial sehingga bukan merupakan persamaan linear dalam $\beta_0, \beta_1, \beta_2, \dots, \beta_p$maka digunakan metode numerik Newton Raphson dalam mencari taksirannya.

Metode Newton-Raphson adalah metode pendekatan yang menggunakan satu titik awal dan mendekati dengan memperhatikan gradien pada titik tersebut. Titik pendekatan ke n+1 dituliskan dengan:

$$\beta^{n+1} = \beta^{(n)} - \frac{l'\left(\beta^{(n)}\right)}{l''\left(\beta^{(n)}\right)}$$

Dengan:

$$\beta = \beta_0, \beta_1, \beta_2, \dots, \beta_p$$

$n = 0,1,2,\dots,k$

$\theta^{(0)}$ = nilai awal yang di pilih

Dalam hal perhitungan iterasi digunakan *software* untuk memperoleh nilai taksiran β.Sehingga taksiran dari model regresi Poisson berganda yaitu

$$\ln\left(\hat{\mu}_i(x_i)\right) = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_p X_{pi}$$

2.2.4 Uji Serentak Parameter Model Regresi Poisson

Menurut Darnah (2011) pengujian serentak parameter model regresi Poisson digunakan untuk mengetahui ada tidaknya pengaruh variabel prediktor terhadap variabel respon.

Hipotesis

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

$H_1 : \exists\beta_j \neq 0$; untuk suatu $j = 1,2, \dots, p$

Taraf Signifikansi

$$\alpha = 0{,}05$$

Statistik uji yang digunakan

$$G = -2\ln\left(\frac{L(\widehat{W})}{L(\widehat{\Omega})}\right) = 2\left[\ln L\left(\widehat{\Omega}\right) - \ln L(\widehat{W})\right] \quad (2.11)$$

Dengan

$L(\widehat{W})$ adalah nilai *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor,

$L(\widehat{\Omega})$ adalah nilai *likelihood* untuk model lengkap dengan melibatkan variabel prediktor.

Kriteria pengujian

Tolak H_0 apabila $G \geq \chi^2_{v,\alpha}$dengan v adalah banyaknya parameter model, artinya ada setidaknya satu variabel prediktor (independen) yang mempengeruhi variabel respon.

2.2.5 Uji Parsial Parameter Regresi Poisson

Menurut Darnah (2011) pengujian secara parsial digunakan untuk mengetahui apakah variabel prediktor berpengaruh terhadap variabel respon secara individual yang dihasilkan. Statistik uji yang digunakan untuk uji parsial yaitu uji *Wald*.

Menurut Listiyani dan Purhadi (2007) hipotesis yang digunakan adalah

$H_0 : \beta_j = 0$; untuk suatu $j = 1,2, \dots, p$

$H_1 : \beta_j \neq 0$; untuk suatu $j = 1,2, \dots, p$

Taraf signifikansi

$$\alpha = 0,05$$

Statistik uji Wald

$$t_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

dengan

$\hat{\beta}_j$ adalah taksiran parameter β_j,

$SE(\hat{\beta}_j)$ adalah taksiran standar eror dari β_j

Kriteria pengujian

H_0ditolak jika $|t_{hit}| > t_{\alpha/2,v}$ atau tolak H_0 jika nilai signifikansi kurang dari α dimana α adalah tingkat signifikansi dan v adalah derajat bebas, artinya variabel prediktor (independen) berpengaruh terhadap variabel respon.

2.2.6 Uji Goodness Of Fit

Menurut Lungan (2006: 267) uji kesesuaian (*goodness of fit*) bertujuan untuk mengambil kesimpulan tentang sebaran populasi. Suatu contoh acak dipilih dari populasi bersangkutan, kemudian informasi contoh tersebut digunakan untuk menguji kebenaran sebaran populasi tersebut. Uji ini didasarkan pada seberapa baik kesesuaian / kecocokan (*goodness of fit*) antara frekuensi pengamatan yang diperoleh data sampel dengan frekuensi harapan yang diperoleh dari distribusi yang dihipotesiskan. Uji *goodness of fit* untuk mengetahui suatu data berdistribusi Poisson atau tidak dengan menggunakan tes *Kolmogorov Smirnov* uji Poisson. Dalam kasus ini akan ditunjukkan bahwa apakah data memiliki distribusi Poisson atau tidak dengan menggunakan software SPSS 16.

Jika dalam perhitungan χ^2 mempunyai nilai yang kecil maka menunjukkan terdapat kecocokan yang baik antara frekuensi harapan dan frekuensi pengamatan sehingga akan terjadi penerimaan H_0 atau penolakan H_1. Sebaliknya, jika dalam perhitungan χ^2 mem-punyai nilai yang besar maka menunjukkan kecocokan yangjelek antara frekuensi harapan dan frekuensi pengamatan sehingga akan terjadi penerimaan H_1atau penolakan H_0 .

Misalkan Y adalah sebuah variabel random yang menyatakan data *count*. Untuk *uji goodness of fit* digunakan sampel berisi n buah pengamatan yang saling bebas dari variabel Y. Pengamatan-pengamatan tersebut kemudian dikelompokkan ke dalam k buah kelas.

Hipotesis :

$H_0 : F(X) = F_0(X)$ (Sampel berasal dari populasi yang berdistribusi Poisson).

$H_1 : F(X) \neq F_0(X)$ (Sampel tidak berasal dari populasi yang berdistribusi Poisson).

Taraf Signifikansi $\alpha = 0{,}05$

Kriteria uji

$$D = \left| F_0(X) - F(X) \right|$$

H_0 ditolak jika nilai $D > D_\alpha$, dimana D_α merupakan nilai kritis yang di peroleh dari tabel *kolmogorov smirnov*. DanH_0 ditolak jika nilaisig hasil output memiliki nilai kurang dari 0,05. artinya sampel tidak berdistribusi poisson.

2.3 Parameter Dispersi

Menurut Darnah (2011) parameter dispersi (ϕ) diperoleh dari rumus

$\phi = \frac{nilai\ deviance}{df}$

dengan

df (degree of freedom) :n-k-1

k : banyaknya parameter termasuk konstanta

n : banyaknya pengamatan

μ_i : penduga bagi respon rata-rata ke-i

Menurut Rashwan dan Kamel (2011) nilai *deviance* didefinisikan sebagai

$$Deviance: G^2 = 2\sum_{i=1}^{n} y_i \ln \frac{y_i}{\mu_i}$$

Apabila nilai $\phi > 0$ maka terjadi *overdispersi* dan apabila $\phi < 0$maka terjadi *underdispersi*.

2.4 *Overdispersi* dan *Underdispersi*

Dalam menganalisis data hasil*count* banyak ditemukan kasus data yang memiliki nilai variansinya lebih besar atau lebih kecil dari nilai rata-ratanya. Untuk menganalisis data diskrit biasanya digunakan regresi Poisson. Namun, dalam regresi Poisson asumsi yang harus dipenuhi adalah adanya *equidispersi* atau nilai variansinya sama dengan nilai rata-ratanya.

Menurut Darnah (2011) *overdispersi*adalah kondisi dimana data variabel respon menunjukkan nilaivariansi lebih besar dari nilai rata-ratanya. *Underdispersi* adalah kondisi dimana data variabel respon menunjukkan nilai variansi lebih kecil dari nilai rata-ratanya.

Overdispersi ataupun underdispersi akan menghasilkan nilai devians model menjadi sangat besar sehingga model yang dihasilkan kurang tepat. Nilai devians diperoleh dari nilai *Deviance* dibagi dengan derajat kebebasan (dilihat pada output SPSS). Salah satu model yang dapat digunakan untuk mengatasi masalah *overdispersi* dan *underdispersi* adalah dengan menggunakan model regresi Poisson tergeneralisasi. Model regresi ini merupakan perluasan dari regresi

Poisson dan baik digunakan dalam keadaan *equidispersi*, *overdispersi*dan *underdispersi*.

Menurut Irwan dan Sari (2013), ketika model Poisson diaplikasikan untuk data *overdispersi*, menyebabkan *standarerror underestimate*. Akibatnya, bebera-pavariabel penjelas menjadi tidak signifikan.

2.5 Multikolinearitas

Multikolinearitas berarti keberadaan dari hubungan linear yang sempurna atau tepat di antara sebagian atau seluruh variabel penjelas dalam sebuah model regresi (Gujarati dan Porter, 2010: 408). Variabel penjelas dalam hal ini yaitu variabel prediktor (X). Konsekuensi jika dalam sebuah model mengandung multikolinearitas adalah variannya akan terus naik atau membesar. Jika variansemakin naik atau membesar maka *standar error*$\beta_1$danβ_2 juga naik atau membesar. Ada beberapa cara untuk mendeteksi ada tidaknya multikolinearitas.

Menurut Priyatno (2013: 60) untuk mendeteksi ada tidaknya multikolinearitas dengan melihat nilai*Tolerance* dan VIF (*Variance Inflation Factor*). Jika nilai *Tolerance* lebih dari 0,1 dan VIF (*Variance Inflation Factor*)kurang dari 10 maka tidak terjadi multikolinearitas.

Langkahnya

Hipotesis

H_0: Model regresi memiliki masalah multikolinieritas

H_1: Model regresi tidak memiliki masalah multikolinieritas

Taraf Signifikansi $\alpha = 0,05$

Statistik uji

$$VIF = \frac{1}{(1 - r_{i,j}{}^2)}$$

$$Tolerance = \frac{1}{VIF_J}(1 - R_j^2)$$

dimana

$r_{i,j}$ adalah koefisien korelasi antara X_i dengan X_j

R_j^2 adalah R^2 pada regresi dari X_j .

Kriteria Uji

Tolak H_0 jika seluruh variabel prediktor memiliki nilai VIF kurang dari 10 dan nilai *Tolerance* lebih dari 0,1. Sebaliknya jika seluruh variabel prediktor memiliki nilai VIF lebih besar 10 dan nilai *Tolerance* kurang dari 0,1 maka H_0diterima. Jika H_0 ditolak, artinya model regressi tidak memiliki masalah multikolinearitas.

Beberapa cara yang dapat digunakan untuk menyembuhkan kasus multikolinearitas diantaranya yaitu

1. Dengan mengeluarkan variabel yang mempunyai nilai VIF tertinggi. Cara ini merupakan cara yang paling sederhana dalam menangani masalah multikolinearitas, tetapi dapat memungkinkan terjadinya kesalahan / bias spesifikasi.
2. Dengan mentransformasi variabel. Transformasi dapat dilakukan dalam bentuk logaritma natural dan bentuk first difference (diferensing 1).
3. Dengan menggabungkan data crosssection dan time series (pooling data), dengan menggunakan metode analisis Bayesian Regression atau dalam kasus khusus Ridge Regression.

2.6 Model Regresi Poisson Tergeneralisasi (Generalized Poisson Regression)

Salah satu asumsi yang harus dipenuhi dalam model regresi Poisson adalah *Equidispersi*, yaitu kondisi dimana nilai rata-rata dan variansi dari variabel respon bernilai sama. Pada prakteknya kadang terjadi pelanggaran asumsi dalam analisis data diskrit berupa *overdispersi* atau *underdispersi* sehingga model regresi Poisson tidak tepat digunakan.

Menurut Melliana (2013) penanganan pelanggaran asumsi *equidispersi* pada model regresi Poisson dapat dikembangkan dengan menggunakan model *Generalized Poisson Regression* (GPR). Model*Generalized Poisson Regression* (GPR) mirip dengan model regresi Poisson tetapi diasumsikan komponen acaknya didistribusikan keumum Poisson. Dengan kata lain model *Generalized PoissonRegression*(GPR) dapat digunakan untuk data diskrit yang mempunyai distribusi Poisson tanpa adanya asumsi *equidispersi*.

Menurut Sadia (2013), dalam *Generalized Poisson Regression* (GPR) fungsi probabilitasbersyarat dari Y_i diberikan nilai $x_{1i}, x_{2i}, ..., x_{pi}$ adalah:

$$f_i\left(y_i|x_{1i}, x_{2i}, \dots, x_{pi}: \mu_i, \alpha\right) = \left(\frac{\mu_i}{1+\alpha\mu_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\alpha y_i}{1+\alpha\mu_i}\right)$$

$$y_i = 0,1, \dots, \infty \qquad (2.12)$$

Dari (2.12) jelas terlihat bahwa fungsi probabilitas dari Regresi Poisson Tergeneralisir merupakan perluasan dari fungsi probabilitas poissondimana pada fungsi probabilitas Regresi Poisson Tergeneralisir terdapat sebuah parameter tambahan yaitu αyang merupakan parameter dispersi.

Rata-rata dan variansi bersyarat dari Y_i diberikan untuk$X_{1i} = x_{1i}, X_{2i} = x_{2i}, \dots, X_{pi} = x_{pi}$Regresi Poisson Tergeneralisir Menurut Famoye (2004) adalah$E\left(Y_i\middle|X_{1i} = x_{1i}, X_{2i} = x_{2i}, \dots, X_{pi} = x_{pi}\right) = \mu_i$

$$Var\left(Y_i\middle|X_{1i} = x_{1i}, X_{2i} = x_{2i}, \dots, X_{pi} = x_{pi}\right) = \mu_i(1+\alpha\mu_i)^2$$

1. Diketahui Y_i memiliki fungsi distribusi probabilitas seperti yang tertera pada Persamaan (2.12), sehingga diperoleh:

$$E(Y_i) = \sum_{y_i=0}^{\infty} y_i \left(\frac{\mu_i}{1+\alpha\mu_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\alpha y_i)}{1+\alpha\mu_i}\right)$$

$$E(Y_i) = \sum_{y_i=1}^{\infty} y_i \left(\frac{\mu_i}{1+\alpha\mu_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\alpha y_i)}{1+\alpha\mu_i}\right)$$

2. Variansi dari $Y_i$dengan menggunakan $Var(Y) = E(Y^2) - [(E(Y)]^2$ Karena $Y_i$memiliki fungsi distribusi probabilitas seperti yang tertera pada Persamaan (2.12), sehingga diperoleh:

$$E(Y_i^2) = \sum_{y_i=0}^{\infty} y_i^2 \left(\frac{\mu_i}{1+\alpha\mu_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\alpha y_i)}{1+\alpha\mu_i}\right)$$

$$E(Y_i^2) = \sum_{y_i=1}^{\infty} y_i^2 \left(\frac{\mu_i}{1+\alpha\mu_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\alpha y_i)}{1+\alpha\mu_i}\right)$$

Misalkan $y = f(t)$ dan $y_0 = f(x_0)$ dimana $f'(x_0) \neq 0$, maka berlaku :

$$f(t) = f(x_0) + \sum_{n=1}^{\infty} \frac{(y-y_0)^2}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \left[f'(x) \left(\frac{x-x_0}{f(x)-y(0)} \right) \right] \right\}_{x-x_0}$$

selanjutnya, misalkan $f(t) = e^{r(t-1)}$, $f(x) = te^{ar(t-1)}$, $y_0 = 0$ dan $x_0 = 0$, maka akan diperoleh:

$$f(t) = f(0) + \sum_{n=1}^{\infty} \frac{te^{ar(t-1)} - 0}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \left[re^{r(x-1)} \left(\frac{x-0}{xe^{ar(x-1)} - 0} \right)^n \right] \right\}_{x=0}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \left[re^{r(x-1)} \left(\frac{x}{xe^{-ar(x-1)}} \right)^n \right] \right\}_{x=0}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \left[re^{r(x-1)} e^{arn(x-1)} \right] \right\}_{x=0}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \left[re^{(1+\alpha n)r(x-1)} \right] \right\}_{x=0}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} r^{n-1} (1+\alpha n)^{n-1} \left\{ \left[re^{(1+\alpha n)r(x-1)} \right] \right\}_{x=0}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} r^{n-1} (1+\alpha n)^{n-1} \left\{ \left[re^{-(1+\alpha n)r)} \right] \right\}$$

$$= e^{-r} + \sum_{n=1}^{\infty} t^n \frac{e^{arn(t-1)}}{n!} r^{n-1} (1+\alpha n)^{n-1} \left\{ \left[re^{-(1+\alpha n)r)} \right] \right\}$$

$$= e^{-r} + \sum_{n=1}^{\infty} \frac{r^n}{n!} (1+\alpha n)^{n-1} e^{-r(1+\alpha nt)} t^n$$

$$= e^{-r} + \sum_{n=1}^{\infty} \frac{r^n}{n!} (1+\alpha n)^{n-1} u_n(t),\ u_n(t) = e^{-r(1+\alpha nt)} t^n. \tag{2.13}$$

Misalkan $f(t) = e^{r(t-1)}$, maka $f(1) = 0$. Apabila disubstitusikan t = 1 pada persamaan (2.15), maka akan diperoleh:

$$\sum_{n=1}^{\infty} \frac{r^n}{n!} (1+\alpha n)^{n-1} e^{-r(1+\alpha nt)} = 1 \tag{2.14}$$

Perhatikan bahwa $u_n'(t) = nt^{n-1}e^{-r(1+\alpha nt)} - \alpha rne^{-r(1+\alpha nt)}t^n$. Karena deret $\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}u_n'(t)$ konvergen uniform di setiap interval dan terbatas, maka diperoleh:

$$f'(t) = \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}u_n'(t)$$

$$= \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}\left(nt^{n-1}e^{-r(1+\alpha nt)} - \alpha rne^{-r(1+\alpha nt)}t^n\right)$$

$$= \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}nt^{n-1}e^{-r(1+\alpha nt)} - \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}\alpha rne^{-r(1+\alpha nt)}t^n$$

Apabila disubstitusikan t = 1, maka akan diperoleh:

$$f'(t) = \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}ne^{-r(1+\alpha nt)} - \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}\alpha rne^{-r(1+\alpha nt)}t^n$$

$$= (1-\alpha r)\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}ne^{-r(1+\alpha nt)}$$

Sedangkan dengan permisalan $f(t) = e^{r(t-1)}$ maka $f'(t) = re^{r(t-1)}$ dan $f'(1) = r$ sehingga diperoleh:

$$r = (1-\alpha r)\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}ne^{-r(1+\alpha nt)}$$

Sehingga diperoleh: $$\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}ne^{-r(1+\alpha nt)} = \frac{r}{(1-\alpha r)} \tag{2.15}$$

Selanjutnya perhatikan turunan kedua dari $u_n(t)$, yaitu:

$u_n''(t) = ne^{-r(1+\alpha nt)}(r^2\alpha^2nt^n - 2\alpha rnt^{n-1} + n - t^{n-2})$ sehingga dipeooleh,

$u_n''(1) = e^{-r(1+\alpha nt)}(r^2\alpha^2n^2 - 2\alpha rn^2 + n^2)$ jadi:

$$f''(1) = \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}u_n''(t)$$

$$f''(1) = \sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}(r^2\alpha^2n^2 - 2\alpha rn^2 + n^2)$$

$$=\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}r^2\alpha^2n^2-\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}2\alpha rn^2+n^2)$$

$$=(1-\alpha r)^2\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}n^2-\frac{r}{1-\alpha r}$$

$$=(1-\alpha r)^2\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}n^2-\frac{r}{1-\alpha r}$$

Misalkang $f(t)=e^{r(t-1)}$ maka $f''(t)=r^2e^{r(t-1)}$ dan $f''(1)=r^2$.

Sehingga diperoleh:

$$r^2=(1-\alpha r)^2\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}n^2-\frac{r}{1-\alpha r}$$

Jadi:

$$r^2+\frac{r}{1-\alpha r}=(1-\alpha r)^2\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}n^2$$

$$\frac{r^2-\alpha r^2+r}{1-\alpha r}=(1-\alpha r)^2\sum_{n=1}^{\infty}\frac{r^n}{n!}(1+\alpha n)^{n-1}e^{-r(1+\alpha nt)}n^2 \qquad (2.16)$$

Hubungan nilai rata-rata dan varians dalam Regresi Poisson Tergeneralisirdapat dikondisikan,jika nilai varians sama dengan nilai rata-rata $E(Y_i)=Var(Y_i)$, maka nilai parameter disperse $\alpha=0$, sehingga fungsi densitas peluang Regresi Poisson Tergeneralisir, akan diturunkan ke regresi poisson dan jika nilai varians lebih besar dari nilai rata-rata $E(Y_i)<Var(Y_i)$, maka nilai parameter disperse $\alpha>0$, sehingga dapat dikatakan pada data terjadi overdispersi dan jika nilai varians lebih kecil dari pada nilai rata-rata $E(Y_i)>Var(Y_i)$, maka nilai parameter disperse α< 0, sehingga pada data terjadi underdispersi.

Dengan mensubtitusikan $\mu_i=\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji}$ ke dalam (2.12) maka diperoleh:

$$f_i(y_i | x_{1i}, x_{2i}, ..., x_{pi} : \mu_i, \alpha) = \left(\frac{\exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}}{1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}} \right)^{y_i} \frac{(1+\alpha y_i)^{y_i - 1}}{y_i !} \exp\left(- \frac{\exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji} (1 + \alpha y_i}{1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}} \right) \tag{2.17}$$

$y_i = 0,1,2,...,n$ sedangkan $\beta = [\beta_0, \beta_1, \beta_2, ..., \beta_p]^T$ menyatakan vektor dari parameterparameter yang tidak diketahui.

2.7 Estimasi Parameter Model Regresi Poisson Tergeneralisir

Dalam penggunaan Regresi Poisson Tergeneralisir, nilai dari parameter-parameter yang tidak diketahui, yaitu $\beta_0, \beta_1, ..., \beta_p, \alpha$ harus ditaksir. Untuk menaksir parameter tersebut digunakan metode *maximum likelihood*.

Jika n pasang pengamatan $\{(X_{1i}, X_{2i}, ..., X_{pi}, Y_i), i = 1,2,3,...,n\}$ diasumsikan saling bebas, maka fungsi likelihood diperoleh dengan mengalikan semua fungsi probabilitas bersyarat dari Y_i diberikan nilai $X_{1i}, X_{2i}, ..., X_{pi}$ pada (2.17), yaitu:

$$\beta^* = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$L(\beta^*) = \prod_{i=1}^{n} f(Y_i | x_{1i}, x_{2i}, ..., x_{pi}; \beta_0, \beta_1, ..., \beta_p, \alpha)$$

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

Untuk mempermudah perhitungan mendapatkan taksiran maksimum likelihood dari parameter $\beta_0, \beta_1, ..., \beta_p$ dan parameter α digunakan fungsi *log likelihood* sebagai berikut:

$$l(\beta^*) = \log L(\beta^*)$$

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

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

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

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

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

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

(2.18)

Untuk mencari taksiran maksimum likelihood fungsi *loglikelihood* $\hat{\beta}_0,\hat{\beta}_1,...,\hat{\beta}_p,\alpha$, pada (2.18) diturunkan secara parsial terhadap parameter $\hat{\beta}_0,\hat{\beta}_1,...,\hat{\beta}_p,\alpha$,α dan disamakan dengan nol. Didapatkan persamaan sebagai berikut:

$$\frac{\partial l(\beta^*)}{\partial \beta_0} =$$

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

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

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

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

$$=\sum_{i=1}^{n}\left\{\left(\frac{\left(1+\alpha\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji})\right)^2 y_i - 1+\alpha\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji}) y_i\alpha\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji})}{\left(1+\alpha\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji})\right)^2}\right) - \left(\frac{(\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji})+\alpha y_i\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji}))}{\left(1+\alpha\exp(\beta_0+\sum_{j=i}^{p}\beta_j x_{ji})\right)^2}\right)\right\}=0$$

$$\frac{\partial l(\beta^*)}{\partial\beta_1}=$$

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

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

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

$$= \sum_{i=1}^{n} \left\{ \left(\frac{ y_i x_{1i} \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right)^2 - y_i x_{1i} \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) }{ \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right)^2 } \right) - \frac{ (1 + \alpha y_i) \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) }{ \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right)^2 } \right\} = 0$$

Dengan cara yang serupa maka akan diperoleh:

$$\frac{\partial l(\beta^*)}{\partial \beta_k} = \sum_{i=1}^{n} \left(\frac{ x_{ki} \left(y_i - \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right) }{ \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right)^2 } \right)$$

$$\frac{\partial l(\beta^*)}{\partial \beta_\alpha} = \sum_{i=1}^{n} \left(\frac{ \left(-y_i \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right) }{ \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right) } + \left(\frac{ y_i (y_i - 1) }{ 1 + \alpha y_i } \right) - \left(\frac{ \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) y_i - \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) }{ \left(1 + \alpha \exp(\beta_0 + \sum_{j=i}^{p} \beta_j x_{ji}) \right) } \right) \right)$$

Karena persamaan-persamaan tersebut tidak linear dalam parameter, sehingga akan dilakukan pendekatan numerik dengan menggunakan metode *Newton Raphson*

2.8 HIV

2.8.1 Pengertian

HIV yang merupakan singkatan dari HUMAN IMMUNODEFICIENCY VIRUS adalah Virus Penyebab AIDS. Virus ini menyerang dan merusak sistem kekebalan tubuh sehingga kita tidak bisa bertahan terhadap penyakit-penyakit yang menyerang tubuh kita.Sistem kekebalan tubuh Rusak atau Lemah mudah terserang penyakit yang ada di sekitar kita seperti TBC, Diare , Sakit kulit, dll.

AIDS yang merupakan kependekan dari *ACQUIRED IMMUNE DEFICIENCY SYNDROME* adalah sindroma menurunnya kekebalan tubuh yang disebabkan oleh HIV. Orang yang mengidap AIDS amat mudah tertular oleh berbagai macam penyakit, karena sistem kekebalan di dalam tubuhnya telah menurun.Sampai sekarang belum ada obat yang dapat menyembuhkan AIDS.

HIV merupakan penyakit menular yang disebabkan oleh infeksi Human Immunodeficiency Virus. Virus ini menyebar melalui cairan tubuh, dan menyerang sistem kekebalan tubuh, khususnya sel CD4 atau yang sering disebut sel T (CDC, 2016). Infeksi tersebut menyebabkan beberapa sel tubuh hancur sehingga penderita mengalami penurunan ketahanan tubuh dan tidak dapat melawan infeksi maupun penyakit lain. Sistem kekebalan dianggap defisien ketika sistem tersebut tidak dapat lagi menjalankan fungsinya memerangi infeksi dan penyakit- penyakit. Orang yang kekebalan tubuhnya defisien (Immunodeficient) menjadi lebih rentan terhadap berbagai ragam infeksi, yang sebagian besar jarang menjangkiti orang yang tidak mengalami defisiensi kekebalan. Penyakit-penyakit yang berkaitan dengan defisiensi kekebalan yang parah dikenal sebagai "infeksi oportunistik" karena infeksiinfeksi tersebut memanfaatkan sistem kekebalan tubuh yang melemah.

Jumlah HIV positif yang ada di masyarakat dapat diketahui melalui layanan Voluntary, Counseling, and Testing (VCT)(Dinkes Jateng, 2013). HIV menular melalui hubungan seksual yang tidak terlindungi (vaginal maupun anal) dan oral sex dengan orang yang terinfeksi, donor darah, dan berbagi jarum suntik yang sudah terkontaminasi. Hubungan seksual dengan penggunaan kondom dapat meminimalisir penyebaran virus HIV. Jumlah kasus baru HIV di Jawa Tengah dari tahun 2012 hingga tahun 2015 cenderung meningkat.

Kasus HIV di Jawa Tengah didominasi kelompok umur seksual aktif. Upaya pemerintah yang dilakukan untuk mengurangi penyakit HIV diantaranya adalah dengan penambahan fasilitas kesehatan dan tenaga kesehatan. Hal ini diharapkanmampu menekan angka kasus HIV. Seseorang mungkin terjangkit penyakit HIV salah satunya dikarenakan status ekonomi. Penderita dengan status ekonomi rendah cenderung mengalami pengobatan yang terbatas. HIV merupakan penyakit yang banyak tertanam dalam ketimpangan sosial ekonomi (Whiteside, 2002). Belenggu kemiskinan membuat seseorang cenderung memilih pekerjaan yang beresiko terjangkit HIV contohnya pekerja seks komersil. Penyakit HIV juga dipengaruhi oleh tingkat pendidikan, hasil survei SDKI tahun 2012 menyatakan bahwa tingkat pengetahuan akan HIV dari setiap orang berbeda menurut pendidikan yang ditamatkan. Seseorang dengan pendidikan SLTA ke atas cenderung lebih banyak tahu mengenai HIV/AIDS dan cara pencegahannya.

2.8.2 Penanganan

Penyakit HIV Stadium I

Biasanya tanpa gejala (dapat limfadenopati generalisata yang persisten PGL), Mampu melawan infeksi dengan baik, pelan-pelan, jumlah CD4 semakin rendah, dan kehidupan sehari-hari tampak normal.

Adapun cara penangananya yaitu dengan pola hidup yang positif dan sehat (olahraga 20 menit setiap hari, makan teratur dan pemeriksaan dokter berkala, skrining IMS, tes Pap, vaksinasi, seks lebih aman.

Penyakit HIV Stadium II

Infeksi ringan lebih sering daripada biasa: ruam, infeksi kulit, seriawan, demam, infeksi saluran pernapasan atas yang kambuhan, umumnya kehilangan berat badan di bawah 10 persen dan dapat meneruskan kehidupan sehari-hari seperti biasa.

Adapun cara penangananya yaitu dengan pola hidup yang positif dan sehat, pemeriksaan, skrining, sekslebih aman, vaksinasi, Pengobatan dini untuk infeksi, Pertimbangkan *profilaksis (kotrimoksazol)*.

Penyakit HIV Stadium III

Infeksi oportunistik (IO) yang lebih parah, mis. pneumonia, meningitis, kandidiasis mulut, diare kronis, demam terus-menerus, TB paru, kehilangan berat badan lebih dari 10 persen dan kesulitan melakukan kegiatan sehari-hari.Adapun cara penangananya yaitu dengan pola hidup yang positif dan sehat, pemeriksaan, skrining, seks lebih aman, vaksinasi, Terapi antiretroviral (ART), pengobatan dini untuk infeksi dan *profilaksis (kotrimoksazol)*.

Penyakit HIV Stadium IV

IO yang lebih parah, mis. PCP, diare parah, limfoma, TB luar paru, tokso, CMV, meningitis kriptokokkus, sarkoma Kaposi, ensefalopati HIV, kandidiasis saluran makan, kehilangan berat badan parah/wasting, sering sakit parah dan terbaring pada tempat tidur.Adapun cara penangananya yaitu dengan Mengobati *IO*,*ART*, Perawatan rumah sakit atau di rumah *Profilaksis (kotrimoksazol)*.