SKRIPSI

PEMODELAN REGRESI CONWAY MAXWELL POISSON UNTUK MENGATASI PELANGGARAN EQUIDISPERSI PADA REGRESI POISSON

Disusun dan diajukan oleh

EVA RIYANTIE

H051171501

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

PEMODELAN REGRESI CONWAY MAXWELL POISSON UNTUK MENGATASI PELANGGARAN EQUIDISPERSI PADA REGRESI POISSON

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

EVA RIYANTIE

H051171501

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

JANUARI 2022

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa Skripsi yang saya buat dengan judul:

Pemodelan Regresi Conway Maxwell Poisson untuk Mengatasi Pelanggaran Equidispersi pada Regresi Poisson

Adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 31 Januari 2022

<u>EVA RIYANTIE</u>

NIM. H051171501

iii

PEMODELAN REGRESI CONWAY MAXWELL POISSON UNTUK MENGATASI PELANGGARAN EQUIDISPERSI PADA REGRESI POISSON

Disetujui Oleh:

Pembimbing Utama,	Pembimbing Pertama,
Andi Kresna Jaya, S.Si., M.Si.	**Dr. Dr. Georgina Maria Tinungki, M.Si.**
NIP. 19731228 200003 1001	**NIP. 19620926 198702 2001**

Ketua Departemen Statistika

Dr. Nurtiti Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2002

Pada Tanggal: 31 Januari 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama	:	Eva Riyantie
NIM	:	H051171501
Program Studi	:	Statistika
Judul Skripsi	:	Pemodelan Regresi Conway Maxwell poisson untuk Mengatasi pelanggaran Equidispersi pada Regresi Poisson

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

		Tanda Tangan
1. Ketua	: Andi Kresna Jaya, S.Si., M.Si.	(…………………)
2. Sekretaris	: Dr. Dr. Georgina Maria Tinungki, M.Si.	(…………………)
3. Anggota	: Dra. Nasrah Sirajang, M.Si.	(…………………)
4. Anggota	: Drs. Raupong, M.Si.	(…………………)

Ditetapkan di: Makassar

Tanggal: 31 Januari 2022

KATA PENGANTAR

Assalamu'alaikum Warohmatullahi Wabarokatuh.

Alhamdulillahi robbil'alamin, Puji syukur kepada **Allah *Subhanahu Wa Ta'ala*** atas segala limpahan rahmat, nikmat, dan hidayah yang diberikan kepada penulis sehingga dapat menyelesaikan penulisan skripsi dengan judul **"Pemodelan Regresi Conway Maxwell Poisson untuk Mengatasi Pelanggaran Equidispersi pada Regresi Poisson"** sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam.

Salam dan sholawat *InsyaAllah* senantiasa tercurah kepada **Nabi Muhammad *Shallallahu'alaihi Wasallam***, sang kekasih tercinta yang telah memberikan petunjuk cinta dan kebenaran dalam kehidupan.

Dalam penyelesaian skripsi ini, penulis telah melewati perjuangan panjang dan pengorbanan yang tidak sedikit. Namun berkat rahmat dan izin-Nya serta dukungan dari berbagai pihak yang turut membantu baik moril maupun material sehingga akhirnya tugas akhir ini dapat terselesaikan. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang setinggi-tingginya dan penghargaan yang tak terhingga kepada Ayahanda **Capt. H. Ansyaruddin Djafar, M.Mar.** dan Ibunda tercinta **Hj. Andi Erniati, S.Pd.i.** yang telah membesarkan dan mendidik penulis dengan penuh kesabaran dan dengan limpahan cinta, kasih sayang, dan doa kepada penulis yang tak pernah habis, untuk kakak tersayang **Rachmat Rizkiadi A., A.Md.Tra.** serta adik tersayang **Nor Azizah Rizkiani A.** dan **Muh. Dzuljawal Ikram A.** yang selalu membantu jika ada kendala selama penulisan dan menjadi penyemangat untuk segera menyelesaikan masa studi penulis.

Ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Ibu Prof. Dr. Dwia Aries Tina Pulubuhu, MA,** selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin,** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.

3. **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si.,** selaku Ketua Departemen Statistika yang telah seperti orang tua sendiri. Segenap dosen pengajar dan staf **Departemen Statistika** yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Bapak Andi Kresna Jaya, S.Si., M.Si.** selaku Pembimbing Utama penulis yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan ditengah kesibukan beliau serta menjadi tempat berkeluh kesah untuk penulis.
5. **Ibu Dr. Dr. Georgina Maria Tinungki, M.Si.** selaku Pembimbing Pertama penulis yang telah meluangkan waktunya ditengah kesibukan untuk memberikan arahan bagi penulis.
6. **Ibu Dra. Nasrah Sirajang, M.Si.** selaku Dosen Penguji sekaligus Penasehat Akademik dan **Bapak Drs. Raupong, M.Si.** selaku Dosen Penguji, terima kasih telah ikhlas meluangkan waktunya ditengah kesibukan untuk memberikan arahan berupa saran dan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
7. Untuk, **Nurul Wahyuni, S.Si., Fakhriyyah Dj. Junus, S.Si., Anisa Haura Salsa F.Y.** dan Kak **Muh. Fadil, S.Si.** yang telah menjadi teman terbaik sejak awal penulisan skripsi dan senantiasa mendengarkan curhatan, memberikan dorongan, semangat, dan motivasi dalam setiap keadaan sehingga penulis bisa mendapatkan lebih banyak pelajaran hidup.
8. Sahabat terbaik, **Putri Eprilian Manda G., S.Kom., Fathimah, S.E., Dwi Susi Kusuma Wati, S.Ak., Nabilah Salsabila, S.Pd., Ailupita Rizkia B., S.Pd., Intania Ayu H., S.Pd.,** dan **Anjani Endiyan, S.Pd.** yang sampai saat ini masih setia mendengarkan keluh kesah penulis.
9. Teman-teman **Statistika 2017**, terima kasih atas kebersamaan, suka, dan duka selama menjalani pendidikan di Departemen Statistika. Penulis senang mengenal kalian semua, terkhusus **Putri Henri, Salsabilah Ramadhani, Indry Angelin, S.Si.,** dan **Gabrielle Jesica T., S.Si.**

10. Keluarga besar **DISKRIT 2017**, terima kasih telah memberikan pelajaran yang berharga dan arti kebersamaan selama ini kepada penulis. Pengalaman yang berharga telah penulis dapatkan dari teman-teman selama berproses.
11. **Keluarga Mahasiswa FMIPA Unhas** terkhusus anggota keluarga **Himatika FMIPA Unhas** dan **Himastat FMIPA Unhas**, terima kasih atas ilmu yang mungkin tidak bisa didapatkan di proses perkuliahan dan telah menjadi keluarga selama penulis kuliah di Universitas Hasanuddin.
12. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih setinggi-tingginya untuk segala dukungan dan partisipasi yang diberikan kepada penulis semoga bernilai ibadah di sisi Allah ***Subhanahu Wa Ta'ala***.

Penulis berharap skripsi ini dapat memberikan tambahan pengetahuan baru bagi para pembelajar statistika. Penulis menyadari bahwa dalam penulisan tugas akhir ini masih banyak terdapat kekurangan. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga dapat bermanfaat bagi pihak-pihak yang berkepentingan. *Aamiin Yaa Rabbal Alamin.*

Makassar, 6 Januari 2022

Eva Riyantie

ABSTRAK

Regresi Poisson merupakan salah satu regresi nonlinier yang digunakan untuk memodelkan hubungan antara variabel respon yang berupa data diskrit dengan variabel prediktor yang berupa data diskrit atau data kontinu. Asumsi yang harus dipenuhi dalam regresi Poisson yaitu equidispersi. Namun, dalam menganalisis data dengan model regresi Poisson dapat terjadi pelanggaran asumsi equidispersi yaitu overdispersi atau underdispersi. Oleh karena itu, untuk mengatasi hal tersebut diperlukan sebuah model yang dapat mengatasi pelanggaran equidispersi pada regresi Poisson yaitu regresi Conway Maxwell Poisson. Pada penelitian ini, regresi Conway Maxwell Poisson diestimasi menggunakan *Maximum Likelihood Estimation* (MLE). Selanjutnya, model regresi Conway Maxwell Poisson diterapkan pada data jumlah kasus tetanus neonatorum di Indonesia tahun 2019 dengan faktor-faktor yang mengurangi risiko terjadinya penyakit tetanus neonatorum diantaranya persentase cakupan imunisasi TT1 pada ibu hamil (X_1), persentase cakupan imunisasi TT2+ pada ibu hamil (X_2), persentase cakupan imunisasi TT5 pada Wanita Usia Subur (WUS) (X_3), persentase cakupan kunjungan antenatal empat kali (K4) (X_4), serta persentase cakupan persalinan di fasilitas pelayanan kesehatan (X_5). Hasil yang diperoleh menunjukkan dengan menggunakan metode MLE dapat mengatasi overdispersi sehingga diperoleh variabel yang berpengaruh signifikan terhadap jumlah kasus tetanus neonatorum di Indonesia tahun 2019 adalah persentase cakupan kunjungan antenatal empat kali (K4) (X_4) dan persentase cakupan persalinan di fasilitas pelayanan kesehatan (X_5).

Kata Kunci: Conway Maxwell Poisson, *Maximum Likelihood Estimation*, Overdispersi, Tetanus Neonatorum.

ABSTRACT

Poisson regression is a method of nonlinear regression that is used to model the relationship between discrete response variables and discrete or continuous predictor variables. Equidispersion is a Poisson regression assumption that must be achieved. However, when using the Poisson regression model to analyze data, there is a possibility of failing the equidispersion assumption, either in overdispersion or underdispersion. To solve this, we need a method named Conway Maxwell Poisson regression that can overcome the failure of equidispersion in Poisson regression. *Maximum Likelihood Estimation* (MLE) was used to estimate Conway Maxwell Poisson regression in this research. Furthermore, the Conway Maxwell Poisson regression model was applied to data on the number of neonatal tetanus cases in Indonesia in 2019 with factors that reduce the risk of neonatal tetanus, including the percentage of TT1 immunization coverage for pregnant women (X_1), percentage of TT2+ immunization coverage for pregnant women (X_2), percentage of TT5 immunization coverage in women of childbearing age (X_3), percentage of coverage of four antenatal visits (K4) (X_4), and percentage of delivery coverage in health care facilities (X_5). The results obtained show that using the MLE method can overcome overdispersion so the variables that have a significant effect on the number of neonatal tetanus cases in Indonesia in 2019 are the percentage of coverage for four antenatal visits (K4) (X_4) and the percentage of delivery coverage in health care facilities (X_5).

Keyword: Conway Maxwell Poisson, *Maximum Likelihood Estimation*, overdispersion, Tetanus Neonatorum.

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah salah satu metode yang digunakan untuk menjelaskan hubungan sebab akibat antara variabel respon dengan variabel prediktor. Analisis regresi dibedakan atas analisis regresi linier dan analisis regresi nonlinier. Analisis regresi linier mempunyai syarat pemenuhan asumsi linieritas dan asumsi data berdistribusi normal. Namun, seringkali ditemukan beberapa data yang tidak memenuhi asumsi yang disyaratkan regresi linier maka dapat diatasi dengan menggunakan analisis regresi nonlinier. Memodelkan data pada analisis regresi nonlinier dapat menggunakan *Generalized Linear Model* (GLM) (Alwi, dkk., 2018).

Analisis regresi Poisson merupakan salah satu analisis regresi nonlinier yang dimodelkan dengan GLM. Analisis ini digunakan ketika variabel respon yang akan dianalisis berupa data diskrit. Asumsi yang harus dipenuhi dalam regresi Poisson adalah nilai variansi sama dengan nilai *mean* pada variabel respon atau disebut equidispersi. Namun, dalam menganalisis data diskrit dengan model regresi Poisson dapat terjadi pelanggaran asumsi yang nilai variansinya lebih besar dari nilai *mean* (overdispersi) atau nilai variansi yang lebih kecil dari nilai *mean* (underdispersi). Regresi Poisson kurang tepat digunakan untuk mengatasi data yang mengalami overdispersi karena nilai *standard error* menjadi *underestimate*, sehingga kesimpulan yang diperoleh menjadi tidak valid (Putra, dkk., 2013). Metode alternatif yang sering dipakai untuk mengatasi pelanggaran equidispersi yaitu regresi Poisson diperumum dan regresi binomial negatif. Hal tersebut pernah dilakukan oleh Ariani (2018) menggunakan regresi binomial negatif untuk mengatasi kasus overdispersi pada data demam berdarah di Provinsi Jawa Tengah dan Hania (2019) memodelkan regresi Poisson diperumum untuk mengatasi overdispersi pada data jumlah kelahiran bayi di Kota Kendari.

Sellers dan Shmueli (2010) memberikan alternatif lain untuk mengatasi pelanggaran equidispersi yaitu dengan menggunakan model regresi Conway Maxwell Poisson (CMP). Proses pembentukan model regresi Conway Maxwell Poisson berdasarkan pada distribusi Conway Maxwell Poisson. Model regresi Conway Maxwell Poisson memiliki dua parameter yaitu parameter regresi dan parameter dispersi. Lord, dkk., (2007) menggunakan model regresi Conway Maxwell Poisson pada data kecelakaan lalu lintas yang mengalami underdispersi. Borle, dkk., (2007) juga menggunakan model regresi Conway Maxwell Poisson pada data dampak partisipasi survei pada perilaku pelanggan yang mengalami overdispersi dan Afri (2017) menggunakan model regresi Conway Maxwell Poisson pada kasus faktor-faktor yang mempengaruhi banyak komplikasi penyakit *Diabetes Mellitus*. Berdasarkan penelitian tersebut, model regresi Conway Maxwell Poisson dapat mengatasi data yang mengalami overdispersi maupun underdispersi.

Pendekatan model regresi Conway Maxwell Poisson akan diaplikasikan pada data jumlah penderita Tetanus Neonatorum (TN), yang mana data tersebut merupakan data jumlahan sehingga tergolong dalam data diskrit. Tetanus Neonatorum adalah penyakit yang disebabkan oleh basil *Clostridium tetani* yang masuk ke dalam tubuh melalui luka. Penyakit ini menginfeksi bayi baru lahir yang berusia kurang dari 28 hari (Kemenkes, 2019).

Berdasarkan uraian dari latar belakang, penulis tertarik untuk memodelkan regresi Conway Maxwell Poisson dengan mengambil judul **"Pemodelan Regresi Conway Maxwell Poisson Untuk Mengatasi Pelanggaran Equidispersi Pada Regresi Poisson"**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan diatas, maka rumusan masalah dalam penelitian ini yaitu bagaimana model regresi Conway Maxwell Poisson (CMP) menggunakan metode *Maximum Likelihood Estimation* (MLE) pada data jumlah kasus Tetanus Neonatorum di Indonesia tahun 2019 yang mengalami overdispersi?

1.3 Batasan Masalah

Pada Penelitian ini akan diaplikasikan pada data jumlah kasus Tetanus Neonatorum di Indonesia tahun 2019 dengan faktor yang mempengaruhi yaitu ibu hamil dan Wanita Usia Subur (WUS). Data jumlah kasus Tetanus Neonatorum pada penelitian ini mengalami overdispersi sehingga akan di estimasi menggunakan *Maximum Likelihood Estimation* (MLE).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah maka tujuan dari penelitian ini yaitu memperoleh model regresi Conway Maxwell Poisson (CMP) menggunakan metode *Maximum Likelihood Estimation* (MLE) pada data jumlah kasus Tetanus Neonatorum di Indonesia tahun 2019 yang mengalami overdispersi.

1.5 Manfaat Penelitian

Manfaat dalam penelitian ini diharapkan dapat menambah pemahaman teoritis dan praktis bagi peneliti dan pembaca tentang model regresi Conway Maxwell Poisson menggunakan metode *Maximum Likelihood Estimation* (MLE) serta pemahaman menganalisis data jumlah kasus Tetanus Neonatorum yang mengalami overdispersi.

BAB II
TINJAUAN PUSTAKA

2.1 *Generalized Linear Model*

Generalized Linear Models (GLM) merupakan perluasan dari model regresi linier klasik yang mana variabel responnya termasuk salah satu anggota keluarga eksponensial (McCullagh dan Neldar, 1989). Menurut Agresti (2002) terdapat tiga komponen utama pada GLM, sebagai berikut:

1. Komponen Acak

   Komponen acak dari *Generalized Linear Models* (GLM) terdiri dari variabel respon Y dari sebuah distribusi keluarga eksponensial. Bentuk fungsi kepadatan peluang dari distribusi keluarga eksponensial adalah sebagai berikut:

   $$f(y;\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right) \tag{2.1}$$

   Parameter θ disebut dengan parameter natural dan parameter ϕ disebut parameter dispersi. Jika ϕ konstan yang diketahui, maka persamaan (2.1) dinyatakan dalam bentuk sebagai berikut:

   $$f(y;\theta) = \exp\big(y\theta - b(\theta) + c(y)\big) \tag{2.2}$$

2. Komponen Sistematik

   Komponen sistematik dari *Generalized Linear Models* (GLM) adalah hubungan dari sebuah vektor prediktor linier untuk menjelaskan variabel-variabel yang berhubungan dalam sebuah model linier

   $$\eta_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} \tag{2.3}$$

   $$\eta_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \;\; dengan\; i = 1,2,\dots,n \tag{2.4}$$

atau dalam bentuk matriks

$$\boldsymbol{\eta} = \boldsymbol{X\beta} \tag{2.5}$$

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

dengan n merupakan banyaknya amatan dan p merupakan banyaknya variabel prediktor.

3. Fungsi Penghubung (*link function*)

   Fungsi penghubung adalah fungsi yang menghubungkan ekspektasi variabel respon Y dengan variabel-variabel prediktor melalui persamaan linier. Misalkan $\mu_i = E(Y_i)$, fungsi yang menghubungkan μ_i dengan prediktor linier (η_i) adalah $g(.)$ sehingga $g(\mu_i) = \eta_i$. Fungsi penghubung $g(.)$ menghubungkan $E(Y_i)$ dengan variabel prediktor dapat dituliskan sebagai berikut:

$$g(\mu_i) = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \ \ dengan\ i = 1,2, \dots, n \tag{2.6}$$

2.2 Distribusi Poisson

Distribusi Poisson merupakan suatu distribusi yang menyatakan peluang banyaknya hasil percobaan yang terjadi yang mana kejadian tergantung pada selang waktu tertentu dengan hasil pengamatan berupa variabel diskrit. Menurut Cameron dan Trivedi (1998), suatu variabel acak Y yang bertipe diskrit akan mengikuti distribusi Poisson dengan fungsi kepadatan peluang sebagai berikut:

$$f(y, \mu) = \frac{e^{-\mu}\mu^y}{y!} \ \ y = 0,1,2, \dots;\ \mu > 0 \tag{2.7}$$

$$f(y, \mu) \ = \exp\left\{\ln\left(\frac{e^{-\mu}\mu^y}{y!}\right)\right\}$$

$$f(y, \mu) \ = \exp\{\ln(e^{-\mu}\mu^y) - \ln(y!)\}$$

$$f(y, \mu) \ = \exp\{(y\ln(\mu) - \mu) - \ln(y!)\} \tag{2.8}$$

dengan membandingkan persamaan (2.1) dengan persamaan (2.8) maka diperoleh sebagai berikut:

$$y = y;\ \theta = \ln(\mu);\ b(\theta) = -\mu;\ a = 1;\ c(y) = -\ln(y!)$$

dengan demikian terbukti bahwa distribusi poisson termasuk dalam keluarga eksponensial. Distribusi Poisson memiliki asumsi $E(Y) = Var(Y) = \mu$ (Margaretha, dkk., 2019).

Uji statistik yang digunakan untuk mengetahui distribusi Poisson pada data yaitu uji *Kolmogorov-Smirnov* dengan hipotesis pengujian sebagai berikut:

H_0: variabel respon berdistribusi Poisson

H_1: variabel respon tidak berdistribusi Poisson

Statistik uji:

$$D_{hitung} = \text{maks}|F_n(Y) - F_0(Y)| \tag{2.9}$$

dengan

$F_0(Y)$ = fungsi distribusi kumulatif yang dihipotesiskan

$F_n(Y)$ = fungsi distribusi kumulatif yang diamati

Kriteria pengujiannya yaitu Tolak H_0 apabila nilai $D_{hitung} > D_{tabel}$ (Daniel, 1989).

2.3 Multikolinieritas

Multikolinieritas adalah adanya hubungan antara variabel prediktor yang satu dengan variabel prediktor yang lain (Aulele, 2012). Permasalahan yang sering muncul pada multikolinieritas yaitu terjadinya korelasi yang cukup tinggi antara variabel-variabel prediktor. Pendeteksian multikolinieritas dapat dilakukan menggunakan nilai *Variance Inflation Factor* (VIF) dengan persamaan sebagai berikut:

$$VIF_j = \frac{1}{1 - R_j^2}\ dengan\ j = 1.2, \dots, p \tag{2.10}$$

dengan R_j^2 adalah nilai koefisien determinasi antara variabel x_j dengan variabel x lainnya. Nilai $VIF \geq 10$ menunjukkan bukti adanya multikolinieritas (Candraningtyas, dkk., 2013).

2.4 Regresi Poisson

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 (Cahyandari, 2014). Model regresi yang terbentuk oleh variabel acak $Y \sim Poisson$ dan rata-rata (μ) yaitu:

$$y_i = \mu_i + \varepsilon_i, \quad i = 1,2,\dots,n \tag{2.11}$$

dengan

$y_i =$ variabel respon pada pengamatan ke-i

$\mu_i =$ rata-rata jumlah kejadian pada waktu tertentu

$\varepsilon_i =$ *error* acak pada pengamatan ke-i

Regresi Poisson menggunakan GLM agar modelnya dapat digunakan dalam pengamatan, yang mana variabel respon pada regresi Poisson tidak mengharuskan berdistribusi normal. Dalam GLM, terdapat sebuah fungsi $g(.)$ yang menghubungkan μ_i dengan prediktor linier (η_i) yang ditunjukkan pada persamaan (2.6). Model regresi Poisson menggunakan fungsi penghubung logaritma natural, karena rata-rata dari variabel responnya akan berbentuk eksponensial dan menjamin bahwa variabel yang diestimasi dari variabel responnya akan bernilai non-negatif. Rata-rata dari variabel respon dan variabel prediktor yang dihubungkan dengan fungsi penghubung tersebut berbentuk sebagai berikut:

$$g(\mu_i) = \ln(\mu_i) = \boldsymbol{x}_i\boldsymbol{\beta} \tag{2.12}$$

$$\mu_i = \exp(\boldsymbol{x}_i\boldsymbol{\beta}) \tag{2.13}$$

Maka fungsi kepadatan peluang distribusi Poisson dapat dituliskan sebagai berikut:

$$f(y_i, \mu_i) = \frac{e^{-\exp(\boldsymbol{x}_i\boldsymbol{\beta})}(\exp(\boldsymbol{x}_i\boldsymbol{\beta}))^{y_i}}{y_i!} \tag{2.14}$$

dengan y_i adalah variabel respon pada pengamatan ke-i, $\exp(\boldsymbol{x}_i\boldsymbol{\beta})$ adalah rata-rata dari distribusi Poisson, $\boldsymbol{x}_i$ adalah vektor yang berukuran $1 \times c$ yang terdiri dari variabel prediktor dan $\boldsymbol{\beta}$ adalah vektor yang berukuran $c \times 1$ yang terdiri dari parameter regresi, dengan $c = p + 1$ (Agresti, 2002). Oleh karena itu, model regresi Poisson dapat dituliskan sebagai berikut:

$$y_i = \exp(\boldsymbol{x}_i\boldsymbol{\beta}) + \varepsilon_i \,, \qquad i = 1,2 \ldots, n \tag{2.15}$$

2.5 Equidispersi

Pada regresi Poisson terdapat asumsi yang harus dipenuhi, yaitu nilai variansi sama dengan nilai rata-rata dari variabel respon atau biasa disebut dengan equidispersi (Myers, dkk., 2010). Namun, menurut Wang dan Famoye (1997) dalam kenyataan di lapangan sering terjadi pelanggaran asumsi tersebut, yaitu nilai variansinya lebih besar dari nilai rata-rata yang dinamakan overdispersi atau nilai variansinya lebih kecil dari nilai rata-rata yang dinamakan underdispersi. Jika terjadi kasus overdispersi pada data, maka regresi Poisson kurang akurat digunakan untuk analisis, karena berdampak pada nilai *standard error* menjadi *underestimate* (lebih kecil dari nilai sesungguhnya), sehingga kesimpulan yang diperoleh menjadi tidak valid. Overdispersi dapat dideteksi menggunakan nilai *Deviance* (D) yang dibagi dengan derajat bebasnya yang dapat dituliskan sebagai berikut (McCullagh & Nelder, 1989):

$$D = 2 \ln \left[\frac{L(y_i, y_i)}{L(\hat{\mu}_i, y_i)} \right] \tag{2.16}$$

$$\phi = \frac{D}{db}$$

Nilai $db = n - c$ dengan n merupakan banyaknya pengamatan dan $c = p + 1$ merupakan banyaknya parameter. Jika nilai statistik uji dispersi lebih besar daripada satu, maka terjadi overdispersi pada variabel respon (Hilbe, 2011).

2.6 Distribusi Conway Maxwell Poisson

Distribusi Conway Maxwell Poisson merupakan pengembangan dari distribusi Poisson yang pertama kali dikenalkan oleh Conway dan Maxwell. Menurut Shmueli, dkk., (2010) distribusi Conway Maxwell Poisson dapat mengatasi data yang mengalami overdispersi atau underdispersi. Bentuk fungsi kepadatan peluang distribusi Conway Maxwell Poisson sebagai berikut:

$$f(y;\mu,\phi)=\frac{\mu^{y}}{(y!)^{\phi}}\frac{1}{Z(\mu;\phi)}, \qquad \mu>0;\phi\geq 0 \tag{2.17}$$

dengan $Z(\mu,\phi)=\sum_{k=0}^{\infty}\frac{\mu^{k}}{(k!)^{\phi}}$ merupakan konstanta normalisasi, parameter μ_i merupakan nilai *mean* dari distribusi Conway Maxwell Poisson dan ϕ merupakan parameter dispersi. Menurut Widjajati dkk.(2015) menyatakan fungsi $Z(\mu;\phi)$ didekati dengan

$$Z(\mu,\phi)\approx\frac{\exp\left(\phi\mu^{\frac{1}{\phi}}\right)}{\mu^{\frac{\phi-1}{2\phi}}(2\pi)^{\frac{\phi-1}{2}}\sqrt{\phi}} \tag{2.18}$$

Distribusi Conway Maxwell Poisson termasuk dalam distribusi keluarga eksponensial sehingga nilai *mean* dan variansi dapat ditunjukkan dengan menggunakan sifat distribusi keluarga eksponensial (Shmueli, dkk., 2005). Berdasarkan persamaan (2.17) diperoleh bentuk seperti persamaan berikut:

$$f(y;\mu,\phi)=\frac{\mu^{y}}{(y!)^{\phi}}\frac{1}{Z(\mu;\phi)}$$

$$f(y;\mu,\phi)=\exp\left(\ln\left(\frac{\mu^{y}}{(y!)^{\phi}}\frac{1}{Z(\mu;\phi)}\right)\right)$$

$$f(y;\mu,\phi)=\exp\left\{\ln(\mu)^{y}-\ln(y!)^{\phi}-\ln Z(\mu;\phi)\right\}$$

$$f(y;\mu,\phi)=\exp\{y\ln(\mu)-\phi\ln(y!)-\ln Z(\mu;\phi)\}$$

$$f(y;\mu,\phi)=\exp\{(y\ln(\mu)-\ln Z(\mu;\phi))-\phi\ln(y!)\} \tag{2.19}$$

dengan membandingkan persamaan (2.1) dengan persamaan (2.19) maka diperoleh sebagai berikut:

$$y=y;\ \theta=\ln\mu;\ a=1;\ b(\theta)=-\ln Z(\mu;\phi);\ c(y)=-\ln(y!)$$

dengan demikian terbukti bahwa distribusi Conway Maxwell Poisson termasuk dalam keluarga eksponensial.

Shmueli, dkk., (2010) juga menyatakan bahwa untuk memperoleh nilai ekspektasi dan variansi distribusi CMP dinyatakan seperti persamaan berikut:

$$E(Y) = \frac{\partial \ln Z(\mu, \phi)}{\partial \ln \mu}$$

$$Var(Y) = \frac{\partial E(Y)}{\partial \ln \mu}$$

Sehingga diperoleh persamaan sebagai berikut:

$$E(Y) = \mu^{\frac{1}{\phi}} - \frac{\phi - 1}{2\phi} \tag{2.20}$$

$$Var(Y) = \frac{\mu^{\frac{1}{\phi}}}{\phi} \tag{2.21}$$

2.7 Regresi Conway Maxwell Poisson

Model regresi Conway Maxwell Poisson merupakan analisis hubungan antara variabel acak respon yang berupa data diskrit berdistribusi Conway Maxwell Poisson dengan satu atau lebih variabel prediktor. Shmueli, dkk., (2010) menyatakan bahwa hubungan antara variabel respon dan variabel prediktor dalam regresi CMP dapat dinyatakan melalui nilai ekspektasi, sehingga model regresi CMP dapat dituliskan sebagai berikut:

$$E(Y_i) = \exp\left(\frac{\boldsymbol{x}_i\boldsymbol{\beta}}{\phi}\right) - \frac{\phi - 1}{2\phi} \tag{2.22}$$

Pada distribusi Conway Maxwell Poisson telah dijelaskan parameter μ_i merupakan nilai ekspektasi dalam distribusi Poisson yang bernilai positif dan $\boldsymbol{x}_i\boldsymbol{\beta}$ bernilai riil. Hubungan nilai ekspektasi dapat dinyatakan pada persamaan (2.13). Selanjutnya, dengan mensubtitusikan persamaan (2.18) dan (2.13) kedalam persamaan (2.17) diperoleh seperti persamaan sebagai berikut:

$$f(y_i; \boldsymbol{\beta}, \phi) = \frac{\exp(\boldsymbol{x}_i\boldsymbol{\beta})^{y_i} \exp(\boldsymbol{x}_i\boldsymbol{\beta})^{\frac{\phi-1}{2\phi}} (2\pi)^{\frac{\phi-1}{2}} \sqrt{\phi}}{(y_i!)^{\phi} \exp\left(\phi \exp\left(\frac{\boldsymbol{x}_i\boldsymbol{\beta}}{\phi}\right)\right)} \tag{2.23}$$

2.8 Metode *Maximum Likelihood Estimation*

Metode *Maximum Likelihood Estimation* (MLE) dapat digunakan untuk menaksir parameter suatu model yang diketahui fungsi probabilitasnya. Misalkan $Y_1, Y_2, \dots, Y_n$ adalah sampel acak dari populasi dengan fungsi probabilitas $f(y_i; \boldsymbol{\beta}, \phi)$. Fungsi *likelihood* didefinisikan sebagai fungsi probabilitas bersama dari $Y_1, Y_2, \dots, Y_n$ yang dianggap sebagai fungsi dari $\boldsymbol{\beta}$ dan ϕ yang dituliskan sebagai $L(\boldsymbol{\beta}, \phi; y_i)$, yaitu:

$$L(\boldsymbol{\beta}, \phi; y_i) = f(y_1, y_2, \dots, y_n; \boldsymbol{\beta}, \phi) \tag{2.24}$$

$$L(\boldsymbol{\beta}, \phi; y_i) = f(y_1; \boldsymbol{\beta}, \phi) \cdot f(y_2; \boldsymbol{\beta}, \phi) \cdots f(y_n; \boldsymbol{\beta}, \phi)$$

$$L(\boldsymbol{\beta}, \phi; y_i) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\beta}, \phi) \tag{2.25}$$

Estimator maksimum *likelihood* $\widehat{\boldsymbol{\beta}}$ dan $\hat{\phi}$ adalah nilai $\boldsymbol{\beta}$ dan ϕ yang memaksimumkan fungsi *likelihood* $L(\boldsymbol{\beta}, \phi; y_i)$. Namun, lebih mudah bekerja dengan logaritma natural dari fungsi *likelihood*, yaitu $l(\boldsymbol{\beta}, \phi; y_i) = \ln L(\boldsymbol{\beta}, \phi; y_i)$. Karena fungsi logaritma natural adalah fungsi dengan monoton naik, maka nilai yang memaksimumkan fungsi $l(\boldsymbol{\beta}, \phi; y_i)$ sama dengan memaksimumkan fungsi $L(\boldsymbol{\beta}, \phi; y_i)$ sehingga dapat dituliskan sebagai berikut:

$$l(\boldsymbol{\beta}, \phi; y_i) = \ln L(\boldsymbol{\beta}, \phi; y_i) \tag{2.26}$$

$$l(\boldsymbol{\beta}, \phi; y_i) = \ln\left(\prod_{i=1}^{n} f(y_i; \boldsymbol{\beta}, \phi)\right) = \sum_{i=1}^{n} \ln f(y_i; \boldsymbol{\beta}, \phi) \tag{2.27}$$

untuk memperoleh nilai $\boldsymbol{\beta}$ dan ϕ dengan memaksimumkan fungsi $l(\boldsymbol{\beta}, \phi; y_i)$, maka $l(\boldsymbol{\beta}, \phi; y_i)$ diturunkan terhadap $\boldsymbol{\beta}$ dan ϕ kemudian disamakan dengan nol seperti persamaan sebagai berikut:

$$l'(\boldsymbol{\beta}, \phi; y_i) = \frac{\partial l(\boldsymbol{\beta}, \phi; y_i)}{\partial \boldsymbol{\beta}} = 0 \tag{2.28}$$

$$l'(\boldsymbol{\beta}, \phi; y_i) = \frac{\partial l(\boldsymbol{\beta}, \phi; y_i)}{\partial \phi} = 0 \tag{2.29}$$

(Hogg, McKean, & Craig, 2013).

2.9 Pengujian Parameter Model Regresi Conway Maxwell Poisson

Setelah mendapatkan model, untuk memeriksa peranan variabel-variabel prediktor dalam model, perlu dilakukan pengujian terhadap parameter model (β_j, dengan $j = 1,2, \dots, p$) untuk mengetahui variabel prediktor yang memiliki pengaruh terhadap model. Pengujian terhadap parameter model dilakukan secara simultan maupun secara parsial.

2.9.1 Pengujian Simultan Parameter Model

Uji simultan digunakan untuk mengetahui signifikansi parameter β terhadap variabel responnya secara keseluruhan. Pengujian parameter model dengan cara simultan dapat menggunakan uji *Likelihood Ratio* dengan statistik uji G. Statistik uji G adalah uji *Likelihood Ratio* yang digunakan untuk menguji signifikansi pengaruh dari semua variabel prediktor secara keseluruhan terhadap variabel respon.

Hipotesis uji simultan untuk model regresi Conway Maxwell Poisson sebagai berikut:

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

(semua variabel prediktor dalam model tidak berpengaruh terhadap variabel respon)

H_1: paling sedikit ada satu $\beta_j \neq 0, j = 1,2, \dots, p$

(paling sedikit ada satu variabel prediktor dalam model yang berpengaruh terhadap variabel respon)

Statistik uji G dinyatakan seperti persamaan (2.30):

$$G = -2ln\left[\frac{L(\beta_0, \phi)}{L(\boldsymbol{\beta}, \phi)}\right] \tag{2.30}$$

dengan $L(\beta_0, \phi)$ adalah fungsi *likelihood* untuk model yang tidak mengandung variabel prediktor dan $L(\boldsymbol{\beta}, \phi)$ adalah fungsi *likelihood* untuk model yang mengandung semua variabel prediktor. Statistik uji G mengikuti sebaran *chi-square* dengan p adalah banyaknya variabel prediktor dapat disimpulkan bahwa minimal terdapat satu variabel prediktor yang mempengaruhi variabel respon. Daerah penolakan H_0 jika $G > \chi^2_{(\alpha,p)}$ yang berarti dapat disimpulkan bahwa paling sedikit ada satu variabel prediktor yang mempengaruhi variabel respon (Hosmer & Lemeshow, 2000).

2.9.2 Pengujian Parsial Satu-Satu

Pengujian parameter regresi secara parsial dilakukan menggunakan statistik uji *Wald*, statistik uji ini sering digunakan untuk menguji signifikansi parameter regresi secara parsial pada masing-masing variabel prediktor. Pengujian parameter secara parsial untuk model Conway Maxwell Poisson sebagai berikut:

Hipotesis yang akan diuji adalah:

$H_0: \beta_j = 0$ (tidak terdapat pengaruh antara variabel prediktor terhadap variabel respon).

$H_1: \beta_j \neq 0$ (terdapat pengaruh antara variabel prediktor terhadap variabel respon).

Statistik uji Wald didefenisikan sebagai berikut:

$$W_j = \left[\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}\right]^2 \tag{2.31}$$

dengan $\hat{\beta}_j$ sebagai penduga dari β_j dan $SE(\hat{\beta}_j)$ sebagai *standard error* atau penduga galat baku β_j. Daerah penolakan H_0, jika $W_j > \chi^2_{\alpha,p}$.

2.10 *Newton Raphson*

Salah satu metode yang paling sering digunakan dalam optimisasi statistika adalah *Newton Raphson* (NR). Metode *Newton Raphson* merupakan salah satu metode untuk menyelesaikan persamaan $f(x) = 0$. Ciri ciri metode NR sebagai berikut:

1. Memerlukan sebuah hampiran awal, dan
2. Memerlukan perhitungan turunan fungsi $f(x)$ dalam setiap iterasi.

Ciri-ciri metode *Newton* menyatakan bahwa hampiran berikutnya diperoleh dengan cara menarik garis singgung kurva $y = f(x)$ pada titik yang mempunyai absis hampiran sebelumnya hingga memotong sumbu x. Titik potong garis singgung tersebut dengan sumbu x merupakan hampiran berikutnya. Proses berlanjut sampai hampiran yang diperoleh mempunyai syarat keakuratan yang ditentukan.

Misalkan g adalah suatu fungsi, bilangan x pada domain g merupakan titik tetap g jika memenuhi $x = g(x)$.

Iterasi

$$x_{t+1} = g(x), \quad t = 0,1,2,3,\ldots$$

disebut iterasi titik tetap.

Misalkan fungsi f mempunyai turunan pertama f'. Barisan $x_0, x_1, x_2, \ldots$ yang diperoleh dari iterasi:

$$x_{t+1} = x_t - \frac{f'(x_t)}{f''(x_t)} \tag{2.32}$$

disebut fungsi iterasi *Newton Raphson*.

2.11 Tetanus Neonatorum

Tetanus neonatorum merupakan penyakit tetanus yang menyerang bayi yang baru lahir. Bayi baru lahir dapat terkena penyakit tetanus apabila Basil *Clostridium* tetani masuk ke dalam tubuhnya melalui luka. Infeksi ini dapat terjadi akibat pemotongan tali pusat dilakukan dengan alat yang tidak steril. Pada negara berkembang masih banyak ditemukan kasus Tetanus Neonatorum (TN), khususnya negara dengan cakupan persalinan oleh tenaga kesehatan yang rendah (Pusat Data dan Informasi Kemenkes RI, 2019).

Penyebab kematian neonatus di dunia salah satunya disebabkan oleh tetanus neonatorum, yaitu secara global hampir sebesar 14% kematian neonatus disebabkan oleh tetanus neonatorum. Tetanus Neonatorum hingga saat ini masih menjadi masalah kesehatan di dunia. Upaya yang dapat dilakukan untuk mencegah tetanus neonatorum dapat dengan melakukan imunisasi Tetanus Toksoid (TT) yang lengkap pada ibu hamil, perawatan persalinan dan pasca persalinan yang bersih (Sari S. N., 2017). Berikut beberapa faktor yang dapat mengurangi risiko terjadinya penyakit Tetanus Neonatorum sebagai berikut (Rusdiana, 2017):

a. Imunisasi Tetanus Toksoid (TT)

1) Wanita Usia Subur (WUS)

Imunisasi Tetanus Toxoid (TT) diberikan kepada Wanita Usia Subur (WUS) yaitu wanita berusia 15-39 tahun. Imunisasi dilakukan sebanyak 5 kali dengan rentang jarak waktu tertentu. Untuk mengukur persentase cakupan imunisasi TT5 pada wanita usia subur didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ WUS\ yang\ melakukan\ imunisasi\ TT5}{seluruh\ jumlah\ WUS} \times 100\%$$

2) Ibu Hamil

Kelompok ibu hamil yang sudah mendapatkan TT2 sampai dengan TT5 dikatakan mendapatkan imunisasi TT2+. Manfaat pemberian imunisasi TT pada ibu hamil yaitu untuk mencegah tetanus bagi ibu dan bayinya yaitu pada proses persalinan yang mana terdapat luka pada rahim maupun pada tali pusat bayi.

a) Cakupan imunisasi TT1

Untuk mengukur persentase cakupan imunisasi TT1 didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT1}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

b) Cakupan imunisasi TT2

Untuk mengukur persentase cakupan imunisasi TT2 didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT2}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

c) Cakupan imunisasi TT3

Untuk mengukur persentase cakupan imunisasi TT3 didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT3}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

d) Cakupan imunisasi TT4

Untuk mengukur persentase cakupan imunisasi TT4 didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT4}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

e) Cakupan imunisasi TT5

Untuk mengukur persentase cakupan imunisasi TT5 didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT5}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

f) Cakupan imunisasi TT2+

Untuk mengukur persentase cakupan imunisasi TT2+ didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ imunisasi\ TT2\ sampai\ TT5}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

b. Kunjungan antenatal empat kali (K4)

Salah satu bentuk pelayanan kesehatan untuk ibu hamil disebut dengan K4. Pelayanan kesehatan ibu hamil harus memenuhi frekuensi minimal di tiap trimester, yaitu minimal satu kali pada trimester pertama (K1) pada usia kehamilan 0-12 minggu, minimal satu kali pada trimester kedua (K2) pada usia kehamilan 12-24 minggu, dan minimal dua kali pada trimester ketiga (K3 dan K4) pada usia kehamilan 24 minggu sampai menjelang persalinan. Standar waktu pelayanan tersebut dianjurkan untuk menjamin perlindungan terhadap ibu hamil dan janin. Untuk mengukur persentase cakupan kunjungan antenatal empat kali (K4) didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ kunjungan\ antenatal\ 4\ kali\ (K4)}{seluruh\ jumlah\ ibu\ hamil} \times 100\%$$

c. Cakupan persalinan di fasilitas pelayann kesehatan

Persalinan di Fasilitas Pelayanan Kesehatan (Fasyankes) yang ditolong oleh tenaga kesehatan terlatih seperti dokter spesialis kebidanan dan kandungan (SpOG), dokter umum, dan bidan merupakan upaya yang dilakukan untuk menurunkan kasus kematian ibu dan bayi. Untuk mengukur persentase cakupan persalinan di fasilitas pelayanan kesehatan didapatkan dengan rumus sebagai berikut:

$$\frac{Jumlah\ ibu\ hamil\ yang\ melakukan\ persalinan\ di\ fasyankes}{seluruh\ jumlah\ ibu\ bersalin} \times 100\%$$