

**PEMODELAN REGRESI *HURDLE* BINOMIAL NEGATIF DAN
HURDLE CONWAY MAXWELL POISSON PADA
DATA PENYAKIT PARKINSON**

**MODELING HURDLE NEGATIVE BINOMIAL AND HURDLE CONWAY
MAXWELL POISSON REGRESSION ON PARKINSON'S DISEASE DATA**



ANISA HAURA SALSA FATIH YUSUF

H062222009



**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**PEMODELAN REGRESI *HURDLE* BINOMIAL NEGATIF DAN
HURDLE CONWAY MAXWELL POISSON PADA
DATA PENYAKIT PARKINSON**

ANISA HAURA SALSA FATIH YUSUF

H062222009



**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR**

2024

**PEMODELAN REGRESI *HURDLE* BINOMIAL NEGATIF DAN
HURDLE CONWAY MAXWELL POISSON PADA
DATA PENYAKIT PARKINSON**

Tesis

sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

ANISA HAURA SALSA FATIH YUSUF

H062222009

kepada

**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

MAKASSAR

2024

TESIS

PEMODELAN REGRESI *HURDLE* BINOMIAL NEGATIF DAN
HURDLE CONWAY MAXWELL POISSON PADA
DATA PENYAKIT PARKINSON

ANISA HAURA SALSA FATIH YUSUF
H062222009

telah dipertahankan di depan Panitia Ujian Magister pada 23 Oktober
2024

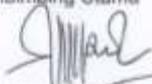
dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Magister Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

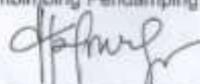
Mengesahkan:

Pembimbing Utama



Prof. Dr. Nurli Sunusi, S.Si., M.Si.
NIP. 19720117 199703 2 002

Pembimbing Pendamping



Dr. Erna Tri Hurdiani, S.Si., M.Si.
NIP. 19750429 200003 2 001

Ketua Program Studi
Magister Statistika


Dr. Erna Tri Hurdiani, S.Si., M.Si.
NIP. 19750429 200003 2 001

Dekan Fakultas Matematika dan
Ilmu Pengetahuan Alam
Universitas Hasanuddin


Dr. Enis Amiruddin, M.Si.
NIP. 19720515 199702 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul "Pemodelan Regresi *Hurdle* Binomial Negatif dan *Hurdle Conway Maxwell Poisson* pada Data Penyakit Parkinson" adalah benar karya saya dengan arahan dari tim pembimbing (Prof. Dr. Nurtiti Sunusi, S.Si., M.Si. dan Dr. Erna Tri Herdani, S.Si., M.Si.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apapun kepada perguruan tinggi manapun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari tesis ini akan dipublikasikan di *Communication in Mathematical Biology and Neuroscience* sebagai artikel dengan judul "*Hurdle Regression Modeling : Parkinson's Disease Data, Specifically Clinical Participants of Movement Disorders Society Unified Parkinson's Disease Rating Scale*".

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 23 Oktober 2024



ANISA HAURA SALSA FATIH YUSUF
NIM. H062222009

UCAPAN TERIMA KASIH

Segala puji hanya milik Allah *Subhanallahu Wa Ta'ala* atas limpahan rahmat dan hidayah-Nya kepada penulis. Shalawat dan salam tercurahkan kepada Rasulullah *Shallallahu 'Alaihi Wa sallam*, keluarganya, *tabi'in, tabi'ut tabi'in*, serta orang-orang sholeh yang haq hingga kadar Allah berlaku atas diri mereka. *Alhamdulillahirobbil'aalamiin*, berkat rahmat dan kemudahan dari Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan tesis berjudul "Pemodelan Regresi *Hurdle Binomial Negatif dan Hurdle Conway Maxwell Poisson* pada Data Penyakit Parkinson" sebagai salah satu syarat memperoleh gelar magister pada Program Studi Magister Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin.

Terima Kasih yang tak terhingga kepada kedua orang tuaku Bapak **Yauri Yusuf** dan Ibu **Evy Santy** yang selalu ada memberikan kasih sayang tak terhingga, semangat, motivasi dan doa yang tak pernah putus, juga kepada adik-adikku **Aura Kaula Yusuf, Muhammad Daffa Fatanah Yusuf, Ariqah Nabilah Zahirah** yang selalu memberikan dukungan dan semangat dalam penyelesaian tesis ini. Ucapan rasa hormat dan juga terima kasih yang tulus kepada:

1. Yth. **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.** selaku Rektor Universitas Hasanuddin.
2. Yth. **Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta seluruhnya jajarannya.
3. Yth. **Prof. Dr. Nurtiti Sunusi, S.Si., M.Si.** selaku Ketua Departemen Statistika sekaligus pembimbing utama penulis yang senantiasa sangat baik dan ikhlas meluangkan waktu dalam memberikan arahan, motivasi, pengetahuan, dan bimbingan untuk penulis dalam menyelesaikan tesis.
4. Yth. **Dr. Erna Tri Herdiani, S.Si., M.Si.** selaku Ketua Program Studi Magister Statistika sekaligus sebagai pembimbing yang telah bersedia menguji serta memberikan masukan-masukan dan arahan dalam penyusunan tesis.
5. Yth. **Dr. Nirwan Ilyas, M.Si.** selaku pembimbing akademik sekaligus penguji yang telah bersedia menguji serta memberikan masukan-masukan dan arahan dalam penyusunan tesis.
6. Yth. **Dr. Anna Islamiyati, M.Si.** selaku penguji yang telah bersedia menguji serta memberikan masukan-masukan dan arahan dalam penyusunan tesis.
7. Yth. **Prof. Dr. Dr. Georgina Maria Tinungki, M.Si.** selaku penguji yang telah bersedia menguji serta memberikan masukan-masukan dan arahan dalam penyusunan tesis.
8. Semua pihak yang telah membantu penulis yang tidak bisa disebutkan satu per satu terimakasih atas doa serta dukungannya.

Semoga Tuhan Yang Maha Esa memberikan balasan yang berlipat ganda, kasih dan hikmat-Nya atas segala kebaikan yang telah diberikan kepada penulis. Penulis menyadari bahwa masih banyak kekurangan dalam tesis ini, untuk itu dengan segala kerendahan hati, penulis memohon maaf.

Makassar, 23 Oktober 2024



Anisa Haura Salsa Fatih Yusuf

ABSTRAK

ANISA HAURA SALSA FATIH YUSUF. **Pemodelan Regresi Hurdle Binomial Negatif dan Hurdle Conway Maxwell Poisson pada Data Penyakit Parkinson** (dibimbing oleh Prof. Dr. Nurtiti Sunusi, S.Si., M.Si. dan Dr. Erna Tri Herdiani, S.Si., M.Si.).

Latar Belakang. Penyakit Parkinson merupakan penyakit yang menyerang bagian motorik pada tubuh sehingga dapat menurunkan kualitas hidup penderitanya. Pengobatan yang dilakukan terbatas dengan pemberian dopaminergik dan terapi fisik. **Tujuan.** Penelitian ini bertujuan untuk mengetahui faktor-faktor yang diduga mempengaruhi komplikasi motorik khususnya pada partisipan klinis *Movement Disorder Society – Unified Parkinson’s Disease Rating Scale* (MDS-UPDRS) yang didiagnosa Penyakit Parkinson. MDS-UPDRS adalah instrumen penilaian yang digunakan untuk mengukur tingkat keparahan gejala komplikasi motorik. Data komplikasi motorik yang diperoleh dari MDS-UPDRS menunjukkan karakteristik distribusi berlebihan pada nilai nol yang berlebih (*zero inflation*), yang mengindikasikan adanya overdispersi akibat *excess zero*. **Metode.** Oleh karena itu, diperlukan regresi *Hurdle* binomial negatif dan *Hurdle Conway Maxwell Poisson* untuk mengatasi hal tersebut. Variabel prediktor yang digunakan merupakan bagian pengukuran dari MDS-UPDRS yaitu aspek motorik, aspek nonmotorik dalam kehidupan sehari-hari dan pemeriksaan motorik oleh profesional medis. **Hasil.** Berdasarkan nilai AIC diperoleh bahwa model regresi *Hurdle* Binomial Negatif lebih baik dibandingkan dengan *Hurdle Conway Maxwell Poisson* untuk diterapkan pada data komplikasi motorik. **Kesimpulan.** Berdasarkan model regresi *Hurdle* binomial negatif diketahui bahwa partisipan klinis yang tidak mengalami komplikasi motorik dipengaruhi secara signifikan oleh variabel aspek nonmotorik dalam kehidupan sehari-hari. Selain itu, pada model *count*, setiap kenaikan skor pada variabel aspek nonmotorik cenderung meningkatkan rata-rata skor komplikasi motorik sebesar 4,1942.

Kata Kunci: penyakit Parkinson; overdispersi; *Hurdle* binomial negatif; *Hurdle Conway Maxwell Poisson*.

ABSTRACT

ANISA HAURA SALSA FATIH YUSUF. **Modeling Hurdle Negative Binomial and Hurdle Conway Maxwell Poisson Regression on Parkinson's Disease Data** (supervised by Prof. Dr. Nurtiti Sunusi, S.Si., M.Si. and Dr. Erna Tri Herdiani, S.Si., M.Si.).

Background. Parkinson's disease is a disease that attacks the motor parts of the body so that it can reduce the quality of life of sufferers. Treatment is limited to dopaminergic and physical therapy. **Aim.** This study aims to determine the factors suspected of influencing motor complications, especially in clinical participants of the Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS) who are diagnosed with Parkinson's disease. MDS-UPDRS is an assessment instrument used to measure the severity of motor complication symptoms. Motor complication data obtained from MDS-UPDRS shows excessive distribution characteristics at excess zero values (zero inflation), which indicates overdispersion due to excess zero. **Method.** Therefore, Hurdle negative binomial and Hurdle Conway Maxwell Poisson regression is needed to overcome this problem. The predictor variables used are part of the measurement of MDS-UPDRS, namely motor aspects, non-motor aspects in daily life, and motor examinations by medical professionals. **Results.** The results shows that Hurdle negative binomial regression model was better than Hurdle Conway Maxwell Poisson when applied to motor complication data. **Conclusion.** Based on the Hurdle negative binomial regression model, it is known that clinical participants who do not experience motor complications are significantly influenced by non-motor aspect variables in daily life. In addition, in the count model, each increase in the score on the non-motor aspect variable tends to increase the average motor complication score by 4.1942.

Keywords: Parkinson's disease; overdispersion, Hurdle negative binomial, Hurdle Conway Maxwell Poisson.

DAFTAR ISI

HALAMAN JUDUL	ii
PERNYATAAN PENGAJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TESIS.....	v
UCAPAN TERIMA KASIH.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	x
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6. Teori.....	3
BAB II METODE PENELITIAN	15
2.1. Sumber Data	15
2.2. Identifikasi Variabel	15
2.3. Metode Analisis Data.....	16
BAB III HASIL DAN PEMBAHASAN.....	17
3.1 Estimasi Parameter Regresi <i>Hurdle</i>	17
3.1.1 Estimasi Parameter Regresi <i>Hurdle</i> Binomial Negatif	17
3.1.2 Estimasi Parameter Regresi <i>Hurdle Conway Maxwell Poisson</i>	21
3.2 Penerapan Data Parkinson pada Model Regresi <i>Hurdle</i> Binomial Negatif dan <i>Hurdle Conway Maxwell Poisson</i>	26
3.2.1 Deskripsi Data Penyakit Parkinson	26
3.2.2 Uji Pelanggaran Equidispersi	28
3.3 Pemodelan Regresi <i>Hurdle</i> Binomial Negatif dan <i>Hurdle Conway Maxwell Poisson</i> pada Data Parkinson.....	29
3.4 Pemilihan Model Terbaik.....	32
BAB IV KESIMPULAN DAN SARAN	33
4.1 Kesimpulan	33
4.2 Saran	34
DAFTAR PUSTAKA.....	35
LAMPIRAN.....	38

DAFTAR TABEL

Nomor urut	Halaman
1. Kriteria Skor Tingkat Keparahan MDS-UPDRS	13
2. Deskripsi Variabel Penelitian	15
3. Frekuensi berdasarkan usia dan jenis kelamin.....	26
4. Tingkat keparahan gejala penyakit Parkinson	27
5. Frekuensi data jumlah skor komplikasi motorik	28
6. Uji overdispersi.....	28
7. Estimasi parameter regresi Hurdle	29
8. Uji likelihood ratio	30
9. Uji Wald regresi Hurdle	31

BAB I

PENDAHULUAN

1.1. Latar Belakang

Menurut *American Association of Neurological Surgeons* (AANS), penyakit Parkinson merupakan gangguan neurodegeneratif yang terjadi karena degenerasi sel saraf di *substantia nigra*, yaitu bagian dari otak tengah yang mengendalikan tubuh. Sel-sel saraf ini mati atau rusak sehingga kehilangan kemampuan untuk memproduksi dopamin. Tanpa dopamin yang cukup, keseimbangan dan koordinasi tubuh dapat terganggu yang dapat menyebabkan tremor, kekakuan otot, dan gerakan yang melambat. Penyakit ini biasanya terjadi pada usia lanjut tetapi individu dengan riwayat penyakit Parkinson dalam keluarga juga memiliki risiko lebih tinggi (AANS, 2024).

Pemberian *dopaminergic* dan terapi fisik merupakan perawatan umum yang saat ini digunakan. Untuk mengevaluasi tingkat keparahan gejala penyakit Parkinson sebelum dan sesudah perawatan, digunakan *Movement Disorder Society Unified Parkinson's Disease Rating Scale* (MDS-UPDRS). Alat ini memungkinkan dokter untuk mendapatkan pemahaman menyeluruh tentang gejala pasien, termasuk gejala motorik dan non-motorik seperti gangguan tidur, masalah *mood* (depresi dan kecemasan), serta disfungsi otonom (gangguan pencernaan dan masalah kardiovaskular) (Cleveland Clinic, 2023).

Namun, beberapa pasien dapat mengalami komplikasi motorik setelah melakukan perawatan. Karakteristik individu yang berisiko tinggi mengalami komplikasi motorik dapat dievaluasi menggunakan tiga bagian pengukuran MDS-UPDRS, bagian I : aspek non-motorik dalam kehidupan sehari-hari, bagian II : aspek motorik dalam kehidupan sehari-hari, dan bagian III : pemeriksaan motorik yang dilakukan langsung oleh profesional medis (Goetz et al., 2008). Regresi Poisson, dapat digunakan untuk mengetahui hubungan antara jumlah komplikasi motorik dengan tiga bagian pengukuran MDS-UPDRS. Asumsi yang harus dipenuhi yaitu equidispersi, nilai rata-rata dan variansi pada data harus sama besar. Namun, terdapat kemungkinan ketika nilai variansi pada variabel respon lebih besar (overdispersi) atau lebih kecil (underdispersi) dari nilai rata-ratanya (Schober & Vetter, 2021).

Ketidaksesuaian data dengan asumsi equidispersi dapat menyebabkan analisis regresi Poisson menghasilkan kesimpulan yang tidak valid (Adams et al., 2020). Oleh karena itu, alternatif untuk menangani salah satu pelanggaran asumsi equidispersi, yaitu overdispersi pada model regresi Poisson adalah menggunakan model regresi binomial negatif. Model ini lebih fleksibel daripada regresi Poisson karena tidak memerlukan pemenuhan asumsi equidispersi yang merupakan kendala dalam regresi Poisson (Schober & Vetter, 2021). Sellers et al. (2012) juga memberikan alternatif lain untuk menangani overdispersi yaitu dengan model regresi

Conway Maxwell Poisson (CMP). Model ini fleksibel dalam menangani data yang mengalami overdispersi atau underdispersi.

Pada beberapa kasus, data *count* yang mengalami overdispersi disebabkan oleh kelebihan nol (*excess zero*). Kelebihan nol terjadi ketika jumlah observasi yang bernilai nol pada variabel respon lebih banyak dari jumlah nilai *count* lainnya (Bekalo & Kebede, 2021). Salah satu metode untuk menangani hal ini yaitu dengan menggunakan model *Hurdle*. Model ini memisahkan data menjadi dua tahap: tahap pertama menggunakan model biner untuk memprediksi kemungkinan nilai nol dan tahap kedua menggunakan model *count* untuk memodelkan data *count* positif. Kelebihan model *Hurdle* adalah kedua model dapat diestimasi secara terpisah sehingga interpretasi model menjadi lebih fleksibel (Bhaktha, 2018). Estimasi parameter regresi *Hurdle* dapat dilakukan dengan algoritma *Broyden-Fletcher-Goldfarb-Shannon* (BFGS). BFGS menggunakan matriks *Hessian* yang dihitung secara iteratif menggunakan turunan pertama dari fungsi sehingga lebih efisien dibandingkan metode *Newton Raphson* yang memerlukan perhitungan turunan kedua untuk matriks *Hessian* (Al-baali *et al.*, 2013).

Adapun penelitian-penelitian yang telah dilakukan untuk mengatasi overdispersi dan kelebihan nol diantaranya: Andika *et al.* (2021) menggunakan model regresi *Hurdle* Binomial Negatif (HBN) pada data *Parkinson's disease*, Sa'diyah *et al.* (2022) memodelkan regresi *Hurdle Poisson* dan HBN pada data jumlah kasus kematian akibat filariasis kronis di Indonesia, Yin *et al.* (2020) melakukan pemodelan regresi *Hurdle Conway Maxwell Poisson* (HCMP) pada data kecelakaan pertambangan, dan Kang *et al.* (2021) melakukan pemodelan HCMP pada data karies gigi di *Lowa School*. Pada penelitian Kang *et al.* (2021) dilakukan perbandingan antara model HCMP dengan *Hurdle Poisson* dan HBN, berdasarkan nilai AIC diperoleh model HCMP lebih baik daripada kedua model lainnya pada data longitudinal dengan studi kasus karies gigi. Namun, pada penelitian Sellers, (2023) dilakukan beberapa pemodelan dengan regresi CMP dan binomial negatif diperoleh berdasarkan nilai AIC bahwa binomial negatif lebih baik daripada CMP untuk mengatasi data yang mengalami overdispersi.

Penelitian ini melanjutkan penelitian Andika *et al.* (2021), dengan menggunakan data penyakit Parkinson dari data klinis MDS-UPDRS (492 pengamatan) yang diperoleh dari *website Parkinson's Progression Markers Initiative* (PPMI, 2024). Analisis awal menunjukkan lebih dari 40% nilai nol pada variabel respon (*excess zero*) dan nilai dispersi data diperoleh lebih besar dari satu (overdispersi). Berdasarkan karakteristik data tersebut, peneliti akan menerapkan dua model regresi pada data penyakit Parkinson, yaitu *Hurdle* binomial negatif dan *Hurdle Conway Maxwell Poisson* yang akan diestimasi menggunakan metode *maximum likelihood estimation*. Oleh karena itu, penulis tertarik melakukan penelitian dengan judul "Pemodelan Regresi *Hurdle* Binomial Negatif dan *Hurdle Conway Maxwell Poisson* pada Data Penyakit Parkinson".

1.2. Rumusan Masalah

Berikut rumusan masalah berdasarkan latar belakang yang telah diuraikan:

1. Bagaimana bentuk estimasi model regresi HBN dan HCMP menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan algoritma *Broyden-Fletcher-Goldfarb-Shannon* (BFGS) pada data penyakit Parkinson?
2. Bagaimana pemodelan data penyakit Parkinson yang mengalami overdispersi menggunakan model HBN dan HCMP?

1.3. Batasan Masalah

Batasan masalah penelitian ini mencakup data klinis MDS-UPDRS yang terdiri dari partisipan yang terdiagnosis penyakit Parkinson dan berada pada tahap awal Parkinson, khususnya partisipan yang melakukan sepuluh kunjungan pemeriksaan dalam waktu lima tahun dengan rentang usia 30 hingga 90 tahun.

1.4. Tujuan Penelitian

Berikut tujuan penelitian berdasarkan rumusan masalah yang telah diuraikan:

1. Memperoleh estimator model HBN dan HCMP menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan algoritma *Broyden-Fletcher-Goldfarb-Shannon* (BFGS) pada data penyakit Parkinson.
2. Memperoleh model HBN dan HCMP pada data penyakit Parkinson yang mengalami overdispersi.

1.5. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Penulis memperoleh pemahaman yang lebih baik mengenai pemodelan regresi HBN dan HCMP pada data *count* yang mengalami masalah overdispersi.
2. Pembaca memperoleh informasi mengenai komplikasi motorik terkait penyakit Parkinson yang mungkin akan mempengaruhi motorik dan nonmotorik di kehidupan sehari-hari.

1.6. Teori

Sub-bab teori ini membahas mengenai beberapa teori yang digunakan untuk membahas pokok pembahasan terkait regresi *Hurdle* binomial negatif dan *Hurdle Conway Maxwell Poisson*.

1.6.1 Generalized linear model

McCullagh & Nelder (1989) menyatakan bahwa *Generalized Linear Model* (GLM) merupakan pengembangan dari model regresi linier klasik ketika tidak terpenuhinya asumsi distribusi normal pada variabel respon. Variabel respon pada GLM diasumsikan mengikuti distribusi dari keluarga eksponensial. Adapun tiga komponen dalam GLM, yaitu (Agresti, 2002):

1. Komponen Acak

Variabel respon $Y = (y_1, y_2, \dots, y_n)$ dalam suatu n pengamatan diasumsikan independen dan mengikuti distribusi keluarga eksponensial, dengan fungsi kepadatan peluang sebagai berikut:

$$f(y_i; \theta_i, \phi) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right) \quad (1)$$

θ_i : parameter natural dengan $i = 1, 2, \dots, n$

ϕ : parameter dispersi

Jika ϕ konstan yang diketahui, maka Persamaan (1) dinyatakan dalam bentuk:

$$f(y_i; \theta_i) = a(\theta_i) b(y_i) \exp[y_i Q(\theta_i)] \quad (2)$$

$$Q(\theta_i) = \frac{\theta_i}{a(\phi)}$$

$$a(\theta_i) = \exp\left[-\frac{b(\theta_i)}{(\phi)}\right]$$

$$b(y_i) = \exp[c(y_i, \phi)]$$

2. Komponen Sistematis

Komponen sistematis merupakan komponen kedua dalam GLM yang menghubungkan vektor η_i dengan variabel prediktor menggunakan prediktor linier. Jika $x_{i1}, x_{i2}, \dots, x_{ip}$ menunjukkan nilai dari sejumlah p variabel prediktor untuk pengamatan ke- i maka η_i dapat dinyatakan dengan persamaan sebagai berikut:

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (3)$$

$$\eta_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}, \text{ dengan } i = 1, 2, \dots, n \quad (4)$$

atau dalam bentuk matriks

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

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

3. Fungsi Penghubung

Fungsi penghubung $g(\cdot)$ adalah suatu fungsi yang menghubungkan komponen acak dengan komponen sistematis. Jika $\mu_i = E(y_i)$, maka model yang menghubungkan μ_i dengan prediktor linier η_i dinyatakan dengan persamaan:

$$g(\mu_i) = \eta_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}, \text{ dengan } i = 1, 2, \dots, n \quad (5)$$

1.6.2 Regresi Poisson

Regresi Poisson adalah metode yang digunakan untuk menganalisis data *count* pada variabel respon yang mengikuti distribusi Poisson. Jika y_i merupakan jumlah kejadian dengan sejumlah n pengamatan dan $\mu_i = E(y_i)$ maka fungsi kepadatan peluang untuk y_i adalah (Agresti, 2002):

$$f(y_i; \mu_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} = \exp(-\mu_i) \exp(y_i \ln \mu_i) \left(\frac{1}{y_i!}\right), y = 0, 1, 2, \dots, n \quad (6)$$

Persamaan (6) memiliki *natural exponential* berdasarkan Persamaan (2) dengan $\theta_i = \mu_i$, $a(\mu_i) = \exp(-\mu_i)$, $b(y_i) = \frac{1}{y_i!}$, dan $Q(\mu_i) = \ln \mu_i$. Parameter natural dari persamaan tersebut adalah $\ln \mu$ dengan fungsi penghubung kanonik *log natural*, $\eta_i = \ln \mu_i$. Jika hubungan μ_i dengan prediktor linier η_i dinyatakan dengan $\ln(\mu_i) = \eta_i$ maka diperoleh model regresi Poisson sebagai berikut:

$$\begin{aligned} \ln \mu_i &= \eta_i \\ \ln \mu_i &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \\ \mu_i &= \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}) \end{aligned} \quad (7)$$

1.6.3 Pelanggaran equidispersi

Regresi Poisson memiliki asumsi utama yaitu equidispersi, nilai variansi dan nilai rata-rata pada variabel respon bernilai sama. Namun, asumsi ini sering dilanggar. Overdispersi (nilai variansi lebih besar dari nilai rata-ratanya) dan underdispersi (nilai variansi lebih kecil dari nilai rata-ratanya) adalah pelanggaran yang umum ditemui. Jika data yang mengalami pelanggaran equidispersi dilanjutkan dengan analisis regresi Poisson maka akan ditemukan estimator yang konsisten tetapi tidak efisien (Hilbe, 2014). Uji *deviance* dilakukan untuk memeriksa pelanggaran equidispersi yang mungkin terjadi dalam suatu data. Jika nilai hitung *deviance* yang dibagi dengan derajat bebas bernilai lebih besar dari satu maka menunjukkan bahwa data mengalami overdispersi. Sebaliknya, jika hasil bagi nilai *deviance* dengan derajat bebasnya kurang dari satu maka menunjukkan bahwa data mengalami underdispersi. Uji *deviance* dihitung dengan persamaan berikut (Agresti, 2002):

$$deviance = 2 \sum_{i=1}^n \left[y_i \log \frac{y_i}{\hat{y}_i} - (y_i - \hat{y}_i) \right] \quad (8)$$

db: derajat bebas, $db = n - (p + 1)$

y_i : nilai variabel respon ke- i dengan $i = 1, 2, \dots, n$

\hat{y}_i : $\exp(X_i \hat{\beta}_{poisson})$

1.6.4 Distribusi binomial negatif

Distribusi binomial negatif adalah pendekatan yang efektif digunakan untuk menangani overdispersi dalam data *count*. Distribusi binomial negatif memiliki parameter dispersi yang mengontrol variabilitas data. Adapun fungsi kepadatan peluangnya adalah sebagai berikut (Bhaktha, 2018):

$$f(y; \mu, \phi) = \frac{\Gamma(y + \phi)}{\Gamma(y + 1)\Gamma(\phi)} \left(\frac{\mu}{\mu + \phi} \right)^y \left(\frac{\phi}{\mu + \phi} \right)^\phi \quad (9)$$

dengan y merupakan nilai data *count*, $\mu = E(y)$, dan ϕ^{-1} merupakan parameter dispersi.

1.6.5 Distribusi Conway Maxwell Poisson

Distribusi *Conway Maxwell Poisson* (CMP) merupakan pengembangan dari distribusi Poisson yang pertama kali dikenalkan oleh Conway dan Maxwell. Menurut Sellers et al. (2012) distribusi *Conway Maxwell Poisson* dapat mengatasi data yang mengalami overdispersi atau underdispersi. Distribusi CMP mengontrol tingkat dispersi data

dengan parameter v , yaitu equidispersi ($v = 1$), overdispersi ($0 \leq v < 1$), dan underdispersi ($v > 1$). Distribusi ini juga mencakup beberapa distribusi umum sebagai kasus khusus: distribusi Poisson ($v = 1$), distribusi *geometric* ($v = 0$ dan $\mu < 1$), dan distribusi Bernoulli ($v \rightarrow \infty$). Berikut bentuk fungsi kepadatan peluang distribusi CMP:

$$f(y; \mu, v) = \frac{\mu^y}{(y!)^v Z(\mu, v)}, \quad \mu > 0; v \geq 0 \quad (10)$$

dengan $Z(\mu, v) = \sum_{y=0}^{\infty} \frac{\mu^y}{(y!)^v}$ merupakan konstanta normalisasi, $\mu = E(y)$, dan v merupakan parameter dispersi (Sellers & Premeaux, 2020).

1.6.6 Regresi *Hurdle*

Salah satu metode untuk mengatasi overdispersi adalah regresi *Hurdle*. Overdispersi terjadi ketika variansi data lebih besar dari nilai rata-ratanya. Salah satu penyebab overdispersi adalah nilai nol yang berlebih pada variabel respon (*excess zero*). Regresi *Hurdle* membagi model menjadi dua bagian yaitu: model untuk data biner dengan nilai nol atau positif yang ditaksir dengan regresi logistik dan model untuk data yang bernilai positif ditaksir dengan *truncated model*. Distribusi probabilitas regresi *Hurdle* adalah sebagai berikut (Mullahy, 1986).

$$f(y_i) = \begin{cases} f_1(0), & y_i = 0 \\ \frac{1 - f_1(0)}{1 - f_2(0)} f_2(y_i), & y_i > 0 \end{cases} \quad (11)$$

$f_1(0)$ adalah probabilitas ketika $y_i = 0$

$f_2(y_i)$ adalah fungsi kepadatan peluang y_i ketika $y_i > 0$

1.6.7 Regresi *Hurdle* binomial negatif

Regresi *Hurdle* Binomial Negatif (HBN) adalah regresi *Hurdle* yang bagian model keduanya menggunakan distribusi binomial negatif. Regresi ini efektif untuk data yang menunjukkan adanya overdispersi akibat nol yang berlebih pada variabel respon (Desjardins, 2013). Jika $y_i = 1, 2, \dots, n$ merupakan variabel respon sebanyak n dari pengukuran observasi ke- i maka fungsi kepadatan peluang dari regresi HBN yang dibentuk berdasarkan Persamaan (11) (Saffari et al., 2012):

$$P(Y_i = y_i) = \begin{cases} (1 - \rho_i), & y_i = 0 \\ (\rho_i) \frac{\Gamma(y_i + \phi)}{\Gamma(y_i + 1)\Gamma(\phi)} \left(\frac{\mu_i}{\phi + \mu_i}\right)^{y_i} \frac{(1 + \phi^{-1}\mu_i)^{-\phi}}{1 - (1 + \phi^{-1}\mu_i)^{-\phi}}, & y_i > 0 \end{cases} \quad (12)$$

Peluang terjadinya nilai selain nol didefinisikan sebagai π_i , dengan $0 < \pi_i < 1$. μ_i adalah rata-rata dari distribusi binomial negatif dengan $\mu_i > 0$ dan ϕ^{-1} adalah parameter dispersi dengan $\phi > 0$. Diketahui bahwa π_i dan μ_i bergantung pada vektor dari variabel prediktor yang dapat didefinisikan sebagai berikut:

a. Model *zero Hurdle* (model logistik)

$$\begin{aligned}\text{logit}(\rho_i) &= \ln\left(\frac{\rho_i}{1-\rho_i}\right) = \sum_{j=1}^p z_{ij}\delta_j \\ \frac{\rho_i}{1-\rho_i} &= \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right) \\ \rho_i &= (1-\rho_i) \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right) \\ \rho_i &= \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right) - (\rho_i) \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right) \\ \rho_i \left(1 + \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right)\right) &= \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right)\end{aligned}$$

maka diperoleh:

$$\rho_i = \frac{\exp\left(\sum_{j=1}^p z_{ij}\delta_j\right)}{1 + \exp\left(\sum_{j=1}^p z_{ij}\delta_j\right)} \quad (13)$$

dengan $0 < \rho_i < 1$, $i = 1, 2, \dots, n$, dan $j = 1, 2, \dots, p$.

Penduga model untuk *zero Hurdle* dengan fungsi penghubung logit dinyatakan sebagai berikut:

$$\text{logit}(\rho_i) = \hat{\delta}_0 + \sum_{j=1}^p z_{ij} \hat{\delta}_j \quad (14)$$

atau dalam bentuk matriks, Persamaan (14) dapat ditulis

$$\begin{bmatrix} \text{logit } \rho_1 \\ \text{logit } \rho_2 \\ \vdots \\ \text{logit } \rho_n \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & \dots & z_{1p} \\ 1 & z_{21} & \dots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{n1} & \dots & z_{np} \end{bmatrix} \begin{bmatrix} \hat{\delta}_0 \\ \hat{\delta}_1 \\ \vdots \\ \hat{\delta}_p \end{bmatrix}$$

b. Model *truncated* binomial negatif

Selanjutnya, nilai μ_i didapatkan dari model *log natural* sebagai berikut:

$$\begin{aligned}\ln(\mu_i) &= \sum_{j=1}^p x_{ij}\beta_j \\ \mu_i &= \exp\left(\sum_{j=1}^p x_{ij}\beta_j\right)\end{aligned} \quad (15)$$

dengan $i = 1, 2, \dots, n$, dan $j = 1, 2, \dots, p$.

Penduga model *truncated* binomial negatif dengan fungsi penghubung *log natural* dinyatakan sebagai berikut:

$$\ln(\mu_i) = \hat{\beta}_0 + \sum_{j=1}^p x_{ij} \hat{\beta}_j \quad (16)$$

atau dalam bentuk matriks, Persamaan (16) dapat ditulis

$$\begin{bmatrix} \ln \mu_1 \\ \ln \mu_2 \\ \vdots \\ \ln \mu_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix}$$

1.6.8 Regresi *Hurdle Conway Maxwell Poisson*

Regresi *Hurdle Conway Maxwell Poisson* (HCMP) menggunakan regresi *Hurdle* dengan model kedua berdasarkan distribusi CMP. Regresi ini menggunakan parameter v untuk menyesuaikan dispersi pada data sehingga fleksibel dalam mengatasi overdispersi dan underdispersi pada data. Jika $y_i = 1, 2, \dots, n$ merupakan variabel respon sebanyak n dari pengukuran observasi ke- i maka fungsi kepadatan peluang dari regresi HCMP berdasarkan Persamaan (11) (Sellers, 2023):

$$P(Y_i = y_i) = \begin{cases} (1 - \rho_i), & y_i = 0 \\ (\rho_i) \frac{1}{Z(\mu_i; v) - 1} \frac{\mu_i^{y_i}}{(y_i!)^v}, & y_i > 0 \end{cases} \quad (17)$$

Sellers (2023) menyatakan fungsi $Z(\mu_i; v)$ dapat didekati dengan

$$Z(\mu_i, v) \approx \frac{\exp\left(v\mu_i^{\frac{1}{v}}\right)}{\mu_i^{\frac{v-1}{2v}} (2\pi)^{\frac{v-1}{2}} \sqrt{v}} \quad (18)$$

a. Model *zero Hurdle* (model logistik)

$$\rho_i = \frac{\exp\left(\sum_{j=1}^p z_{ij} \delta_j\right)}{1 + \exp\left(\sum_{j=1}^p z_{ij} \delta_j\right)}$$

dengan $0 < \rho_i < 1$, $i = 1, 2, \dots, n$, dan $j = 1, 2, \dots, p$.

Penduga model untuk *zero Hurdle* dengan fungsi penghubung logit dinyatakan sebagai berikut:

$$\text{logit}(\rho_i) = \hat{\delta}_0 + \sum_{j=1}^p z_{ij} \hat{\delta}_j$$

atau dalam bentuk matriks dapat ditulis

$$\begin{bmatrix} \text{logit } \rho_1 \\ \text{logit } \rho_2 \\ \vdots \\ \text{logit } \rho_n \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & \dots & z_{1p} \\ 1 & z_{21} & \dots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{n1} & \dots & z_{np} \end{bmatrix} \begin{bmatrix} \hat{\delta}_0 \\ \hat{\delta}_1 \\ \vdots \\ \hat{\delta}_p \end{bmatrix}$$

b. Model *truncated* CMP

Selanjutnya, nilai μ didapatkan dari model *log natural* sebagai berikut:

$$\ln(\mu_i) = \sum_{j=1}^p x_{ij} \gamma_j \quad (19)$$

$$\mu_i = \exp\left(\sum_{j=1}^p x_{ij}\gamma_j\right)$$

dengan $i = 1, 2, \dots, n$, dan $j = 1, 2, \dots, p$.

Penduga model *truncated* CMP dengan fungsi penghubung *log natural* dinyatakan sebagai berikut:

$$\ln(\mu_i) = \hat{\gamma}_0 + \sum_{j=1}^p x_{ij} \hat{\gamma}_j \quad (20)$$

atau dalam bentuk matriks dapat ditulis

$$\begin{bmatrix} \ln \mu_1 \\ \ln \mu_2 \\ \vdots \\ \ln \mu_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_p \end{bmatrix}$$

1.6.9 Pengujian parameter

1. Uji simultan

Uji *likelihood ratio* atau uji G_{hit} digunakan untuk mengetahui variabel prediktor yang secara simultan memiliki pengaruh signifikan terhadap variabel respon.

a. Hipotesis

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_p = 0$$

$$H_1 : \text{Ada } \theta_j \neq 0, j = 1, 2, \dots, p$$

b. Statistik Uji

$$G_{hit} = -2 \ln \left[\frac{L_0(\hat{\theta})}{L_1(\hat{\theta})} \right] = -2 [\ln L_0(\hat{\theta}) - \ln L_1(\hat{\theta})] \quad (21)$$

L_0 : nilai fungsi *likelihood* untuk model tanpa variabel prediktor

L_1 : nilai fungsi *likelihood* untuk model dengan semua variabel prediktor.

c. Kriteria Pengujian

Jika $G_{hit} > \chi^2_{(\alpha, p)}$ maka tolak H_0 .

d. Kesimpulan

Penolakan H_0 menunjukkan bahwa variabel prediktor secara simultan memiliki pengaruh yang signifikan terhadap variabel respon.

2. Uji parsial

Uji *Wald* digunakan untuk mengetahui variabel prediktor yang secara parsial memiliki pengaruh signifikan terhadap variabel respon.

a. Hipotesis

$$H_0 : \theta_j = 0$$

$$H_1 : \theta_j \neq 0, j = 1, 2, \dots, p$$

b. Statistik Uji

$$W(\hat{\theta}_j) = \left(\frac{\hat{\theta}_j}{SE(\hat{\theta}_j)} \right)^2 \quad (22)$$

$\hat{\theta}_j$: estimator dari θ_j
 $SE(\hat{\theta}_j)$: *standard error* dari estimator $\hat{\theta}_j$

c. Kriteria Pengujian

Jika $W_j > \chi^2_{(\alpha,1)}$ maka tolak H_0 .

d. Kesimpulan

Penolakan H_0 menunjukkan bahwa variabel prediktor ke- j memiliki pengaruh yang signifikan terhadap variabel respon.

1.6.10 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 $f(y_i; \theta)$. Fungsi *likelihood* didefinisikan sebagai fungsi probabilitas bersama serangkaian data pengamatan Y_1, Y_2, \dots, Y_n yang merupakan fungsi dari parameter θ yang dituliskan sebagai $L(\theta; y_i)$, yaitu:

$$\begin{aligned} L(\theta; y_i) &= f(y_1, y_2, \dots, y_n; \theta) \\ L(\theta; y_i) &= f(y_1; \theta) \cdot f(y_2; \theta) \cdots f(y_n; \theta) \\ L(\theta; y_i) &= \prod_{i=1}^n f(y_i; \theta) \end{aligned} \quad (23)$$

Estimator *maximum likelihood* $\hat{\theta}$ adalah nilai θ yang memaksimumkan fungsi *likelihood* $L(\theta; y_i)$. Namun, lebih mudah bekerja dengan logaritma natural dari fungsi *likelihood*, yaitu $l(\theta; y_i) = \ln L(\theta; y_i)$. Karena fungsi logaritma natural adalah fungsi yang monoton naik, maka nilai yang memaksimumkan fungsi $l(\theta; y_i)$ sama dengan memaksimumkan fungsi $L(\theta; y_i)$ sehingga dapat dituliskan sebagai berikut (Hogg et al., 2019):

$$\begin{aligned} l(\theta; y_i) &= \ln L(\theta; y_i) \\ l(\theta; y_i) &= \ln \left[\prod_{i=1}^n f(y_i; \theta) \right] = \sum_{i=1}^n \ln f(y_i; \theta) \end{aligned} \quad (24)$$

untuk memperoleh nilai θ yang memaksimumkan fungsi $l(\theta; y_i)$, maka $l(\theta; y_i)$ diturunkan terhadap θ kemudian disamakan dengan nol seperti persamaan berikut:

$$l'(\theta; y_i) = \frac{\partial l(\theta; y_i)}{\partial \theta} = 0 \quad (25)$$

1.6.11 Algoritma *Broyden-Fletcher-Goldfarb-Shanno*

Algoritma *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) merupakan pengembangan dari metode *Newton Raphson*. Perbedaan utamanya terletak pada perhitungan matriks *Hessian*. Pada *Newton Raphson*, matrik *Hessian* dihitung langsung dari turunan kedua fungsi *log-likelihood*. Sedangkan pada BFGS, matriks *Hessian* dihitung secara iteratif menggunakan informasi yang diperoleh dari setiap langkah iterasi. Estimasi parameter diperoleh dengan persamaan berikut (Ullah et al., 2021):

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \lambda_k \mathbf{H}_k^{-1} \mathbf{g}_k \quad (26)$$

dengan

$$\lambda_k = \min \ln L(\boldsymbol{\theta}_k - \lambda \mathbf{H}_k^{-1} \mathbf{g}_k)$$

\mathbf{g} : turunan pertama dari fungsi *log likelihood* ($\nabla f(\boldsymbol{\theta}_k)$).

\mathbf{H} : matriks Hessian

$$\mathbf{H}_{k+1} = \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{c}_k^T}{\mathbf{c}_k^T \mathbf{s}_k} \right) \mathbf{H}_k \left(\mathbf{I} - \frac{\mathbf{c}_k \mathbf{s}_k^T}{\mathbf{c}_k^T \mathbf{s}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{c}_k^T \mathbf{s}_k} \quad (27)$$

$$\mathbf{s}_k = \boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k$$

$$\mathbf{c}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$$

\mathbf{H}_0 : matriks identitas $n \times n$

Proses iterasi akan berlanjut hingga estimasi parameter konvergen yaitu pada saat $\|\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k\| \leq \varepsilon$, dengan ε adalah bilangan yang sangat kecil mendekati nol.

1.6.12 Pemilihan model terbaik

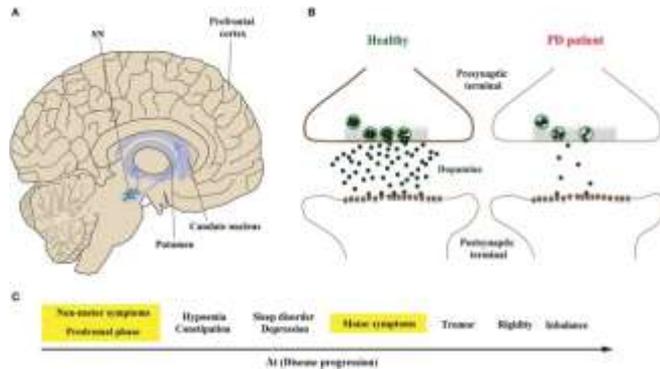
Salah satu alat penting untuk dalam pemilihan model terbaik adalah dengan menggunakan *Akaike Information Criterion* (AIC). AIC menilai seberapa baik model mewakili data dengan mempertimbangkan jumlah parameter dan tingkat kecocokan dengan data. Model dengan nilai AIC yang lebih rendah dianggap lebih baik. Perbedaan AIC antara dua model mengindikasikan seberapa besar salah satu model lebih baik daripada yang lain. Nilai AIC dapat didefinisikan:

$$AIC = -2 \log L(\hat{\boldsymbol{\theta}}) + 2p \quad (28)$$

dengan $L(\hat{\boldsymbol{\theta}})$ adalah nilai *likelihood* dari model yang diberikan data dan p adalah jumlah parameter dalam model (Emiliano et al., 2014).

1.6.13 Penyakit Parkinson

Penyakit Parkinson, gangguan *degenerative* otak pada *Substantia Nigra*, berkembang perlahan dengan gejala yang muncul bertahun-tahun. Penderita biasanya hidup lama dengan kehilangan kontrol motorik secara progresif yang disebabkan oleh defisit hormon dopamin, disertai depresi, gangguan lambung, dan gangguan kognitif lainnya. Gambar 1. menunjukkan *substantia nigra*, yaitu bagian otak tengah yang mengandung neuron penghasil dopamin.



Gambar 1. Degenerasi *Substantia Nigra*

Sumber : (Bridi & Hirth, 2018)

Gejala utama akibat defisit hormon dopamin meliputi tremor, gerakan yang lambat, dan otot yang kaku. Meskipun penyebab penyakit Parkinson tidak diketahui, beberapa faktor dianggap berperan, seperti genetik atau terpapar racun di lingkungan (Mahadewa & Ngoerah, 2022). Menurut *World Health Organization* (WHO) penyakit ini umumnya diderita oleh lanjut usia, tetapi juga dapat diderita oleh orang yang lebih muda. Jika dibandingkan dengan perempuan, laki-laki lebih rentan (WHO, 2023).

WHO (2022) menyatakan bahwa dalam dua puluh lima tahun terakhir, prevalensi penyakit Parkinson telah meningkat dua kali lipat. Lebih dari 8,5 juta orang diperkirakan menderita Parkinson pada tahun 2019. Sejak tahun 2000, penyakit ini telah meningkatkan jumlah *disability adjusted life years* (DALYs) sebesar 81%, mencapai 5,8 juta, dan jumlah kematian meningkat lebih dari 100%, mencapai 329.000 kematian. Hingga kini, belum ada obat yang ditemukan untuk menyembuhkan penyakit Parkinson, tetapi yang tersedia hanya mampu meredakan gejala pada pasien (WHO, 2023). Terapi yang dapat diberikan untuk mengurangi gejala diantaranya memberikan obat dopaminergik seperti levodopa, latihan fisik, terapi bicara, antidepresan, dan toksin botulinum untuk gangguan motorik. Tujuan dari terapi ini adalah untuk meningkatkan kualitas hidup penderita (Mahadewa & Ngoerah, 2022). Alat yang digunakan untuk mengukur tingkat keparahan dan perkembangan gejala penyakit Parkinson sebelum dan sesudah diberikan perawatan yaitu menggunakan *Movement Disorder Society Unified Parkinson's Disease Rating Scale* (MDS-UPDRS).

Bagian-bagian pengukuran dari MDS-UPDRS, yaitu bagian I : aspek non-motorik dalam kehidupan sehari-hari, bagian II : aspek motorik dalam kehidupan sehari-hari, bagian III : pemeriksaan motorik oleh profesional medis, dan bagian IV : komplikasi motorik. MDS-UPDRS merupakan revisi dari *Unified Parkinson's Disease Rating Scale* (UPDRS) yang awalnya dikembangkan pada tahun 1980an. MDS-UPDRS dikembangkan untuk mengevaluasi berbagai aspek penyakit Parkinson. Hal ini mencakup evaluasi motorik dan mengkarakterisasi tingkat dan beban penyakit di

berbagai populasi. Skala ini dapat digunakan dalam lingkungan klinis maupun penelitian (Goetz et al., 2008). Berikut tabel kriteria skor untuk masing-masing bagian MDS-UPDRS ditunjukkan pada Tabel 1:

Tabel 1. Kriteria Skor Tingkat Keparahan MDS-UPDRS

Bagian MDS-UPDRS	Kategori Keparahan	Rentang Skor (s)
Bagian I : non-motorik dalam kehidupan sehari-hari	Ringan	$s \leq 10$
	Sedang	$11 \leq s \leq 21$
	Berat	$s > 21$
Bagian II : motorik dalam kehidupan sehari-hari	Ringan	$s \leq 12$
	Sedang	$13 \leq s \leq 29$
	Berat	$s > 29$
Bagian III : pemeriksaan motorik oleh profesional medis	Ringan	$s \leq 32$
	Sedang	$33 \leq s \leq 58$
	Berat	$s > 58$
Bagian IV : komplikasi motorik	Ringan	$s \leq 4$
	Sedang	$5 \leq s \leq 12$
	Berat	$s > 12$

Sumber : (Martínez-Martín et al., 2015)

Pada tahun 2010, *The Michael J. Fox Foundation* dan sekelompok peneliti akademis serta mitra industri membuat Parkinson's Progression Markers Initiative (PPMI) untuk menemukan gejala biologis yang signifikan dan mengetahui onset serta perkembangan penyakit Parkinson. Sejak saat itu, PPMI telah melibatkan ribuan peneliti akademis, profesional medis, pakar industri, filantropis, relawan studi untuk membangun landasan penelitian penyakit Parkinson. Analisis dari kumpulan data dan pustaka biosampel yang tersedia dapat diakses di website <https://ida.loni.usc.edu/pages/access/studyData.jsp?project=PPMI>. PPMI terdiri dari tiga jenis data diantaranya (PPMI, 2024):

a. PPMI *Clinical*

Data ini berisi data statis partisipan klinis, riwayat klinis melalui kuisisioner MDS-UPDRS, penilaian non-motorik seperti Montreal Cognitive Assessments, analisis biospesimen, uji genetik, hasil MRI dan DaTscan, serta riwayat klinis.

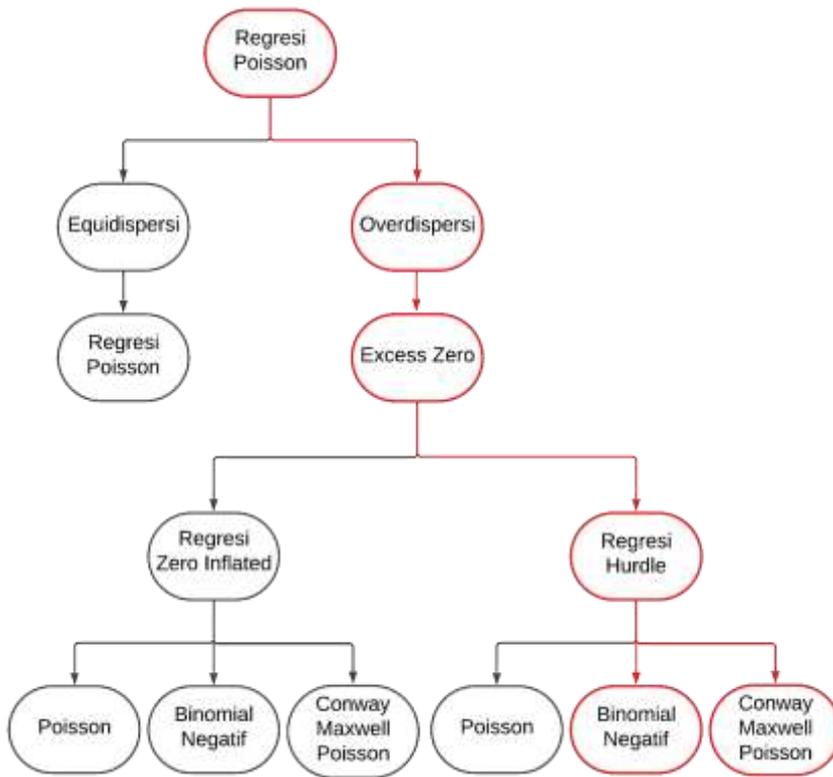
b. PPMI *Remote*

Data dikumpulkan melalui *remote* (jarak jauh) yang berfokus pada *pre-diagnostic phase* penyakit Parkinson. Data ini meliputi gejala olfaktori dan uji genetik (dalam beberapa kasus) untuk mengidentifikasi gejala awal penyakit Parkinson.

c. PPMI *Online*

Penginputan data di lakukan oleh partisipan klinis secara online melalui *website* PPMI. Data ini mencakup sebagian dari data yang ada di PPMI *Clinical*, seperti hanya bagian I dan II dari kuisisioner MDS-UPDRS yang dapat diisi secara mandiri, sedangkan di PPMI *clinical*, bagian I-IV dapat diisi secara lengkap. Selain itu, terdapat data dari partisipan dengan dan tanpa diagnosis Parkinson yang telah di konfirmasi oleh dokter.

1.6.14 Kerangka Konseptual



Gambar 2. Kerangka Konseptual

BAB II METODE PENELITIAN

2.1. Sumber Data

Data yang digunakan merupakan data sekunder yang diperoleh dari *Parkinson's Progression Markers Initiative* (PPMI) versi update 17 April 2024, diakses 18 April 2024 (<https://ida.loni.usc.edu/pages/access/studyData.jsp?project=PPMI>). Jumlah pengamatannya terdiri dari 362 partisipan klinis yang terkait PPMI. Data sekunder yang dimaksud berupa data klinis *Movement Disorder Society – Unified Parkinson's Disease Rating Scale* (MDS-UPDRS) yaitu jumlah skor bagian I : aspek non-motorik dalam kehidupan sehari-hari, jumlah skor bagian II : aspek motorik dalam kehidupan sehari-hari, jumlah skor bagian III : pemeriksaan motorik oleh profesional medis, dan jumlah skor bagian IV : komplikasi motorik yang ditunjukkan pada Lampiran 1.

2.2. Identifikasi Variabel

Variabel yang digunakan dalam penelitian merupakan bagian pengukuran dari MDS-UPDRS yang menggunakan skala Likert dengan rentang nilai skor 0 – 4, skor yang lebih tinggi menunjukkan dampak gejala yang lebih berat. Deskripsi masing-masing variabel dapat dilihat pada Tabel 2.

Tabel 2. Deskripsi Variabel Penelitian

Variabel	Deskripsi
Jumlah skor komplikasi motorik (Y_i)	Data jumlah komplikasi motorik partisipan klinis yang melakukan sepuluh kunjungan dalam 5 tahun (V12). Terdiri dari 6 item, yang merujuk pada komplikasi terkait dengan penggunaan terapi pengobatan Parkinson. Digunakan untuk menggambarkan gejala atau kondisi yang muncul sebagai hasil dari fluktuasi respons terhadap obat.
Jumlah Skor dari bagian I : aspek non-motorik dalam kehidupan sehari-hari (X_1)	Terdiri dari 13 item, yang mencakup berbagai aspek seperti masalah tidur, kognitif, <i>mood</i> , perubahan perilaku, dan gejala <i>autonome</i> . Evaluasi ini membantu dokter untuk memahami secara holistik gejala non-motor yang mempengaruhi kualitas hidup pasien.
Jumlah Skor dari bagian II : aspek motorik dalam kehidupan sehari-hari (X_2)	Terdiri dari 13 item, bagian ini menilai kemampuan melakukan berbagai aktivitas seperti makan, mandi, berjalan, berkomunikasi, dan melakukan kegiatan rumah tangga. Evaluasi ini memberikan gambaran tentang sejauh mana gejala motorik membatasi kemandirian pasien dalam kehidupan sehari-hari.
Jumlah Skor dari bagian III: pemeriksaan motorik oleh profesional medis (X_3)	Terdiri dari 33 item, bagian ini mencakup penilaian terhadap berbagai aspek gejala motorik seperti tremor, kekakuan, bradikinesia, postur, dan juga menilai simetri atau asimetri gejala antara sisi kanan dan kiri tubuh. Evaluasi ini untuk membantu dokter dalam memantau perkembangan penyakit, merancang pengobatan yang sesuai, dan menilai dampaknya terhadap kualitas hidup pasien.

2.3. Metode Analisis Data

Adapun langkah-langkah yang dilakukan berdasarkan tujuan penelitian adalah sebagai berikut:

1. Melakukan analisis deskriptif data untuk memberikan gambaran mengenai karakteristik data.
2. Melakukan pengujian dispersi dengan uji *deviance* menggunakan Persamaan (8).
3. Melakukan pemeriksaan kelebihan nol (*excess zeros*) pada variabel respon berdasarkan proporsi nilai nol yang lebih besar dari nilai positif lainnya.
4. Melakukan estimasi parameter model regresi HBN dan HCMP menggunakan metode MLE.
 - a. Membentuk fungsi *likelihood* kemudian menentukan turunan pertama dari fungsi \ln *likelihood*-nya.
 - b. Jika hasil persamaan yang diperoleh dari turunan pertama fungsi \ln *likelihood* terhadap masing-masing parameter tidak dapat diperoleh secara analitik, maka dilakukan pendekatan algoritma BFGS.
5. Menggunakan algoritma BFGS sebagai metode optimasi untuk menentukan estimasi parameter.
6. Mendapatkan model penduga regresi HBN dan HCMP dari data penyakit Parkinson yang mengalami overdispersi berdasarkan hasil estimasi parameter yang didapatkan.
7. Melakukan uji signifikansi parameter secara simultan dengan uji *likelihood ratio*.
8. Melakukan uji signifikansi parameter secara parsial untuk masing-masing parameter model HBN dan HCMP dengan uji *wald*.
9. Melakukan pemilihan model terbaik berdasarkan nilai AIC.