PEMODELAN REGRESI NONPARAMETRIK POISSON DENGAN ESTIMATOR *PENALIZED SPLINE* PADA KASUS JUMLAH KEMATIAN NEONATUS DI SULAWESI SELATAN

SKRIPSI

AHMAD MUKHLISH MURSIDIN

H051201011

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

MEI 2024

PEMODELAN REGRESI NONPARAMETRIK POISSON DENGAN ESTIMATOR *PENALIZED SPLINE* PADA KASUS JUMLAH KEMATIAN NEONATUS DI SULAWESI SELATAN

SKRIPSI

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

AHMAD MUKHLISH MURSIDI

H051201011

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
MEI 2024

LEMBAR PERNYATAAN KEOTENTIKAN

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

Pemodelan Regresi Nonparametrik Poisson dengan Estimator *Penalized Spline* pada Kasus Jumlah Kematian Neonatus di Sulawesi Selatan

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

Makassar, 13 Mei 2024

SEPULUH RIBU RUPIAH 10000 METERAI TEMPEL
CFAF6ALX194995859

Ahmad Mukhlish Mursidin
NIM H051201011

PEMODELAN REGRESI NONPARAMETRIK POISSON DENGAN ESTIMATOR *PENALIZED SPLINE* PADA KASUS JUMLAH KEMATIAN NEONATUS DI SULAWESI SELATAN

Disetujui Oleh:

Pembimbing Utama

Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002

Ketua Program Studi

Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002

Pada 13 Mei 2024

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Ahmad Mukhlish Mursidin

NIM : H051201011

Program Studi : Statistika

Judul Skripsi : Pemodelan Regresi Nonparametrik Poisson dengan Estimator *Penalized Spline* pada Kasus Jumlah Kematian Neonatus di Sulawesi Selatan

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

DEWAN PENGUJI

1. Ketua : Dr. Anna Islamiyati, S.Si., M.Si. (..............................)
2. Anggota : Andi Kresna Jaya, S.Si., M.Si. (..............................)
3. Anggota : Anisa, S.Si., M.Si. (..............................)

Ditetapkan di : Makassar

Tanggal : 13 Mei 2024

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur kehadirat Allah *Subhanahu Wa Ta'ala* atas segala rahmat, hidayah, dan karunia-Nya yang telah dilimpahkan sehingga penulis mampu menyelesaikan penyusunan skripsi ini. Shalawat serta salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya. *Alhamdulillahi robbil'alamin*, berkat nikmat kemudahan dan pertolongan yang diberikan oleh Allah *Subhanahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi dengan judul "**Pemodelan Regresi Nonparametrik Poisson dengan Estimator *Penalized Spline* pada Kasus Jumlah Kematian Neonatus di Sulawesi Selatan**" yang disusun sebagai salah satu syarat akademik untuk memperoleh gelar sarjana pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyelesaian skripsi ini tidak lepas dari bantuan dan motivasi dari berbagai pihak yang senantiasa turut mendukung dalam bentuk moril maupun materil. Sehingga dengan segala keterbatasan kemampuan dan pengetahuan penulis, skripsi ini dapat terselesaikan. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang tulus serta penghargaan yang tinggi kepada Abba **Mursidin *Rahimahullah*** dan Ummi **Arvyaty** yang senantiasa memberikan dukungan penuh, limpahan cinta dan kasih sayang, serta dengan ikhlas menemani penulis dengan doa dan restu mulianya. Penulis takkan pernah sanggup membalas semua kebaikan kalian. Namun, ketahuilah bahwa penulis sangat mencintai kalian dan selalu berusaha untuk membuat kalian bangga. Ucapan terima kasih juga penulis haturkan kepada **Abba Puang**, **Ummi Puang**, serta saudari-saudari dari Ummi, **Kholah Zahrah**, **Mama Pati**, **Unda Sanga**, dan **Unda Nasyrah** sudah menjadi sosok orang tua kedua penulis yang selalu mendoakan, mendukung, dan memberikan bantuan selama penulis merantau di Makassar. Terima kasih juga kepada saudara-saudara tercinta **Mu'arif**, **Mu'afif**, [illegible] dan **Mundzir** yang selalu menghibur dan memberi semangat kepada [illegible] Serta untuk keluarga besar penulis yang tidak dapat disebutkan [illegible]a, penulis ucapkan terima kasih atas doa dan dukungannya.

Penghargaan yang tinggi dan ucapan terima kasih dengan penuh ketulusan dan keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, 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. Anna Islamiyati, S.Si., M.Si.**, selaku Ketua Departemen Statistika sekaligus Dosen Pembimbing Utama penulis yang telah meluangkan waktu dan pemikirannya untuk senantiasa memberikan ilmu, arahan, motivasi, dan afirmasi positif kepada penulis dari awal hingga penyusunan skripsi selesai.
4. **Bapak Andi Kresna Jaya, S.Si., M.Si.** selaku Dosen Penguji sekaligus Penasehat Akademik penulis dan **Ibu Anisa, S.Si., M.Si.** selaku Dosen Penguji yang telah meluangkan waktu dalam memberikan saran serta kritikan yang membangun kepada penulis dalam penyempurnaan skripsi ini.
5. Segenap **Dosen Pengajar** dan **Staf** yang telah memberikan ilmu dan kemudahan kepada penulis dalam berbagai hal selama menempuh pendidikan sarjana di Departemen Statistika.
6. Sahabat-sahabat tercinta, **Fadlan**, **Fahmi**, **Ryan**, dan **Sabil** yang setia menemani perjuangan suka dan duka dalam menyelesaikan skripsi serta mengurus dokumen-dokumen pendukungnya. Terima kasih atas semua momen-momen kebersamaan belajar, diskusi, curhat, dan merayakan setiap pencapaian kecil yang telah dilewati. Kita bagaikan tim yang saling mendukung dan membantu satu sama lain untuk mencapai tujuan bersama.
7. Sahabat-sahabat **RANDOM**, **Aisyah**, **Alif**, **Bahar**, **Dzaky**, **Fadlan**, **Fahmi**, **Krisna**, **Laurine**, dan **Rais** yang menemani penulis dari awal perkuliahan yang bermula dari grup belajar bersama hingga terbentuk grup **Random** yang juga menjadi grup belajar bersama sekaligus menjadi tempat berbagi cerita serta bermain bersama. Penulis merasa sangat bersyukur dipertemukan kalian, terima kasih telah mengukir kenangan manis bersama penulis selama kuliah.

bat-sahabat **A(MONG)**, **Afika**, **Aliah**, **Aulia**, **Ayu**, **Jihan**, **Nadia**, **la**, **Rahmi**, **Theo**, **Yoel**, dan kawan-kawan. Terima kasih untuk rsamaan di labstat ketika berbagi cerita, belajar serta bermain bersama.

9. Keluarga besar **Himastat FMIPA Unhas** khususnya untuk sahabat-sahabat **POIS20N**, **Alisha**, **Ayuni**, **Azalia**, **Azhar**, **Divia**, **Fauziah**, **Hakam**, **Isra**, **Izzul**, **Kur**, **Linda**, **Liza**, **Maryana**, **Nahdia**, **Razy**, **Reza**, **Ryan**, **Sabil**, dan kawan-kawan. Terima kasih untuk segala cerita, pengalaman, serta kenangan suka dan duka selama berproses di himpunan yang tidak didapatkan penulis di proses belajar mengajar di perkuliahan.
10. Teman-teman seperjuangan **STAT'20**. Masa-masa penyusunan skripsi memang penuh dengan lika-liku. Kita semua pasti pernah merasakan lelah, stres, frustasi, dan ingin menyerah. Namun berkat dukungan dan semangat dari kalian, penulis dapat melewati itu dengan lebih mudah. Terima kasih atas segala ilmu, kebersamaan, serta cerita suka dan duka selama kuliah. Penulis pasti akan merindukan momen kebersamaan bersama kalian semua.
11. Teman-teman **KKN Tematik Indeks Pembangunan Manusia Takalar Gelombang 110** khususnya sahabat-sahabat di **Posko Desa Kampung Beru**, terima kasih untuk seluruh pengalaman baru, cerita suka dan duka, serta kenangan kekeluargaannya selama kurang lebih 45 hari di Desa Kampung Beru di Takalar. Terima kasih juga karena walaupun KKN telah selesai, kita masih ada komunikasi, saling dukung, ikut pelatihan, dan bermain bersama.
12. Sahabat-sahabat **OLDO MKS**, **Acid**, **Ainun**, **Anandi**, **Depi**, **Hjr**, **Fkr**, **Kanmel**, **Mapud**, **Pipo**, dan **Tegar** yang sama-sama merantau dari Kendari ke Makassar. Terima kasih karena senantiasa memberikan semangat dan dukungan, mengajak bergaul, serta menemani perjuangan penulis. Dan terima kasih juga kepada sahabat-sahabat **XII MIA OLIMPIADE 2 (OLDO)**.
13. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih untuk segala dukungan, partisipasi, dan apresiasi yang telah diberikan.

Penulis menyadari bahwa masih banyak kekurangan dalam penyusunan skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini dapat memberikan manfaat untuk berbagai pihak.

Makassar, 13 Mei 2024

Ahmad Mukhlish Mursidin

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademika Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: Ahmad Mukhlish Mursidin
NIM	: H051201011
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

"Pemodelan Regresi Nonparametrik Poisson dengan Estimator *Penalized Spline* pada Kasus Jumlah Kematian Neonatus di Sulawesi Selatan"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar tanggal 13 Mei 2024.

Yang menyatakan,

(Ahmad Mukhlish Mursidin)

ABSTRAK

Jumlah kematian neonatus di Provinsi Sulawesi Selatan merupakan data diskrit dan mengikuti pola distribusi Poisson sehingga dapat dimodelkan dengan menggunakan regresi Poisson. Regresi poisson hakikatnya mengikuti model regresi parametrik, namun tidak semua data dapat diasumsikan mengikuti model regresi parametrik karena sering kali asumsi-asumsinya tidak terpenuhi, sehingga model regresi dikembangkan lagi menjadi regresi nonparametrik yang dapat digunakan pada data yang polanya tidak mengikuti model parametrik. Penelitian ini bertujuan untuk mengestimasi jumlah kematian neonatus menggunakan metode regresi nonparametrik Poisson dengan estimator *penalized spline*. Penggunaan estimator *penalized spline* agar memperoleh hasil estimasi yang akurat dengan mempertimbangkan titik knot dan parameter penghalus secara bersamaan. Hasil estimasi menggunakan regresi nonparametrik Poisson dengan estimator *penalized spline* menunjukkan bahwa jumlah kematian neonatus mempunyai dua jenis kurva estimasi optimal yang ditinjau berdasarkan nilai GCV minimum, yaitu model dengan orde 1 dan 3 titik knot serta orde 1 dan 1 titik knot. Berdasarkan hasil estimasi, diperoleh model terbaik yang menunjukkan bahwa Persentase Ibu Hamil yang Mendapatkan TTD, Jumlah Bayi dengan BBLR, Persentase Bayi Baru Lahir yang Mendapatkan IMD, Jumlah Cakupan Imunisasi Td2+ pada Ibu Hamil, dan Jumlah Cakupan Komplikasi Neonatus yang Ditangani merupakan faktor-faktor yang berpengaruh secara signifikan terhadap Jumlah Kematian Neonatus.

Kata Kunci: Neonatus, Poisson, *Penalized Spline*, Titik Knot, Parameter Penghalus.

ABSTRACT

The number of neonates deaths in South Sulawesi Province is discrete data and follows a Poisson distribution pattern so it can be modeled using Poisson regression. Poisson regression essentially follows a parametric regression model, but not all data can be assumed to follow a parametric regression model because the assumptions are often not met, so the regression model is further developed into non-parametric regression which can be used on data whose patterns do not follow the parametric model. This study aims to estimate the number of neonates deaths using the nonparametric Poisson regression method with a penalized spline estimator. The use of a penalized spline estimator to obtain accurate estimation results by considering knot points and smoothing parameters simultaneously. The estimation results using nonparametric Poisson regression with a penalized spline estimator show that the number of neonates deaths has two types of optimal estimation curves which are viewed based on the minimum GCV value, namely models with orders of 1 & 3 knot points and orders of 1 & 1 knot point. Based on the estimation results, the best model was obtained which shows that the number of babies with low birth weight, the percentage of newborns who receive early initiation of breastfeeding, the number of Td2+ immunization coverage in pregnant women, and the number of neonates complications handled are factors that have a significant effect on the number of neonates deaths.

***Keywords*:** *Neonates, Poisson, Penalized Spline, Knot Points, Smoothing Parameters.*

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kematian bayi dominan terjadi pada saat periode neonatal atau periode bulan pertama setelah bayi lahir. Bayi baru lahir periode bulan pertama dikenal dengan istilah neonatus (Sembiring, 2019). Neonatus sangat mudah terdampak risiko gangguan kesehatan yang dapat mengakibatkan kematian. Pada tahun 2020, terdapat sekitar 2,4 juta kematian neonatus. WHO (2020) menetapkan beberapa negara yang memiliki angka kematian neonatal yang tinggi, Indonesia menempati urutan ke 7 di dunia. Salah satu provinsi penyumbang angka kematian neonatal tertinggi di Indonesia yaitu Sulawesi Selatan dengan jumlah kematian neonatus sebanyak 602 kasus pada tahun 2020 (Dinas Kesehatan Sulawesi Selatan, 2021). Perlu adanya tindakan dari pemerintah setempat untuk mengatasi permasalahan tersebut dengan cara memaksimalkan kebijakan-kebijakan terkait di bidang kesehatan dalam mencegah faktor-faktor yang mengakibatkan jumlah kematian neonatus semakin tinggi. Untuk mengetahui faktor-faktor yang berpengaruh signifikan pada jumlah kematian neonatus dapat menggunakan analisis regresi sebagai dasar pemerintah dalam menentukan kebijakan (Abdy, 2019).

Analisis regresi merupakan penerapan ilmu statistika yang digunakan untuk menyelidiki korelasi antara dua atau lebih variabel (Kurniawan, 2016). Salah satu metode analisis regresi yang dapat memodelkan data jumlah kematian neonatus yang berupa data diskrit adalah regresi Poisson (Umami dkk., 2013). Regresi Poisson digunakan untuk menganalisis hubungan antara variabel respon yang berdistribusi Poisson dan diasumsikan mengalami kondisi equidispersi atau nilai ekspektasi dan variansinya bernilai sama (Irwan dkk., 2021). Regresi poisson hakikatnya mengikuti model regresi parametrik yang diasumsikan fungsi regresi dan bentuk pola sudah diketahui, seperti linier, kuadratik, kubik maupun bentuk lainnya. Namun, tidak semua data dapat diasumsikan mengikuti model regresi [illegible]k karena sering kali asumsi-asumsinya tidak terpenuhi sehingga model [illegible]kembangkan lagi agar dapat digunakan untuk data yang belum diketahui [illegible]ngsi regresinya yaitu regresi nonparametrik (Islamiyati dkk., 2022).

Regresi nonparametrik memiliki fleksibilitas yang tinggi dalam memodelkan karena data diharapkan mencari sendiri bentuk estimasinya tanpa dipengaruhi oleh faktor tertentu (Eubank, 1999). Dalam regresi nonparametrik, terdapat beberapa estimator salah satunya yaitu *spline*. *Spline* adalah suatu metode dalam analisis regresi nonparametrik yang merupakan potongan-potongan polinomial yang memiliki sifat tersegmen dan dapat digunakan pada suatu pendekatan ke arah pencocokan data dengan tetap memperhitungkan kemulusan kurva (Sanusi dkk., 2019). Model pada *spline* cenderung mencari sendiri hasil estimasi data karena terdapat titik knot yang merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku data secara signifikan (Nisa dkk., 2022).

Jao (2022) menggunakan Regresi Poisson dalam mengestimasi model data jumlah kematian bayi namun data yang digunakan mengikuti pola nonparametrik sehingga *spline truncated* digunakan sebagai estimator pada model. *Spline truncated* merupakan satu dari beberapa estimator dalam pendekatan model regresi nonparametrik *spline*. Penaksir model regresi nonparametrik *spline* lainnya digunakan pada penelitian Islamiyati (2019) yaitu *Penalized Spline*. *Penalized Spline* merupakan salah satu estimator dari bentuk regresi nonparametrik *spline* yang memiliki titik knot terletak pada titik-titik kuantil dari nilai *unique* variabel prediktor (Wangsih dkk., 2022). Keunggulan dari *Penalized Spline* yaitu dapat memperhitungkan titik knot dan juga parameter penghalus dalam model secara simultan sehingga menghasilkan ketepatan dan kehalusan kurva estimasi (Islamiyati dkk, 2017). Berdasarkan uraian tersebut, penulis menggunakan metode Regresi Nonparametrik Poisson dengan estimator *Penalized Spline* dalam memodelkan faktor-faktor yang memengaruhi kesehatan ibu hamil dan bayi karena memiliki korelasi terhadap jumlah kematian neonatus di Sulawesi Selatan.

1.2 Rumusan Masalah

Berdasarkan latar belakang, berikut masalah yang dapat dirumuskan:

1. Bagaimana bentuk estimasi model Regresi Nonparametrik Poisson dengan estimator *Penalized Spline*?
2. Bagaimana model terbaik dalam memodelkan data jumlah kematian neonatus di Sulawesi Selatan menggunakan Regresi Nonparametrik Poisson dengan estimator *Penalized Spline*?

1.3 Batasan Masalah

Batasan masalah diperlukan agar pembahasan masalah dalam penelitian tidak terlalu luas. Batasan masalah untuk penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan data sekunder jumlah kematian neonatus di provinsi Sulawesi Selatan tahun 2020 yang terbagi menjadi 24 observasi berdasarkan kabupaten/kota.
2. Equidispersi pada data yang berdistribusi Poisson diasumsikan terpenuhi.
3. Algoritma *Full Search* digunakan untuk menentukan orde, jumlah knot, dan parameter penghalus optimal berdasarkan nilai GCV minimum.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, penelitian ini memiliki tujuan sebagai berikut:

1. Memperoleh bentuk estimasi model Regresi Nonparametrik Poisson dengan estimator *Penalized Spline*.
2. Memperoleh model terbaik dalam memodelkan data jumlah kematian neonatus di Sulawesi Selatan menggunakan Regresi Nonparametrik Poisson dengan estimator *Penalized Spline*.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini, yaitu:

1. Menambah pengetahuan dalam mengimplementasikan estimasi model Regresi Nonparametrik Poisson dengan estimator *Penalized Spline* pada data jumlah kematian neonatus di Sulawesi Selatan sehingga dapat menjadi salah satu referensi untuk penelitian selanjutnya.
2. Memberikan model korelasi antara faktor-faktor yang memengaruhi kesehatan ibu hamil dan bayi terhadap jumlah kematian neonatus di Sulawesi Selatan.
3. Memberikan gagasan objektif dari sudut pandang statistika kepada tenaga kesehatan dan pemerintah dalam memberi solusi untuk kasus kematian neonatus di Sulawesi Selatan.

BAB II
TINJAUAN PUSTAKA

2.1 Regresi Nonparametrik

Dalam regresi nonparametrik, tidak ada asumsi tentang bentuk fungsi regresi sehingga memberikan fleksibilitas di dalam bentuk yang mungkin dari fungsi regresi. Terdapat banyak teknik untuk mengestimasi fungsi regresi nonparametrik, antara lain estimator *spline*, *kernel*, *spline*, *wavelet*, dan lain-lain. Model regresi nonparametrik dinyatakan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i, i = 1,2, \dots, n \tag{2.1}$$

dengan y_i adalah variabel respon, $f(x_i)$ dinyatakan sebagai fungsi regresi yang belum diketahui bentuknya, dan ε_i adalah faktor *error* (Eubank, 1999).

Regresi nonparametrik memiliki keunggulan dibandingkan regresi parametrik, yaitu tingkat fleksibilitas yang lebih tinggi sehingga data mampu mencari bentuk estimasi kurva regresi secara mandiri tanpa dipengaruhi faktor dari peneliti. Estimasi $f(x_i)$ pada regresi nonparametrik sama seperti estimasi koefisien-koefisien regresi parametrik sehingga memudahkan dalam menduga fungsi regresi yang lebih umum (Jao dkk., 2022). Stone (1985) mengembangkan model regresi nonparametrik pada Persamaan (2.1) untuk multiprediktor dengan model didefinisikan sebagai berikut:

$$y_i = \sum_{j=1}^{m} f_j(x_{ji}) + \varepsilon_i \tag{2.2}$$

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

2.2 *Generalized Linear Model*

Generalized Linear Model (GLM) merupakan perluasan dari model regresi ketika variabel respon mengikuti distribusi keluarga eksponensial, yaitu normal, poisson, geometrik, binomial, binomial negatif, dan gamma. Salah satu komponen dalam GLM yaitu fungsi penghubung untuk menghubungkan ekspektasi dari espon dengan variabel-variabel prediktor melalui persamaan linier.

ie (2017) menjelaskan bahwa GLM menghubungkan nilai harapan dari respon dengan variabel-variabel prediktor melalui fungsi penghubung sehingga memungkinkan distribusi dari variabel respon berasal dari keluarga distribusi eksponensial. Model GLM dapat dinyatakan sebagai berikut:

$$\eta_i = \beta_0 + \sum_{j=1}^{m} \beta_i x_{ji} \tag{2.3}$$

McCullagh (1989) menyatakan bahwa terdapat komponen acak dalam GLM yang merupakan suatu fungsi untuk mengidentifikasi distribusi dari variabel respon yang berasal dari keluarga eksponensial. Menurut Dobson (2002), suatu variabel acak Y dapat dikatakan anggota dari distribusi keluarga eksponensial jika memiliki model fungsi probabilitas sebagai berikut:

$$f(y) = s(y)t(\theta)\exp\big(a(y)b(\theta)\big)$$

$$f(y) = \exp\big(a(y)b(\theta) + c(\theta) + d(y)\big) \tag{2.4}$$

dengan $a(y), b(\theta), s(y) = \exp d(y)$, dan $t(\theta) = \exp c(\theta)$ merupakan fungsi yang sudah diketahui dan θ merupakan parameter yang memuat β dan μ.

2.3 Regresi Poisson

Regresi Poisson merupakan salah satu bentuk analisis regresi yang digunakan untuk memodelkan jumlah kejadian pada periode waktu dan wilayah tertentu dengan variabel respon yang memiliki jenis data diskrit (*count data*) dan berdistribusi Poisson. Untuk mengetahui data yang diamati berdistribusi Poisson atau tidak, dapat menggunakan uji *Kolmogorov-Smirnov* dengan statistik uji sebagai berikut:

$$D_{hitung} = maks|F_n(y) - F_0(y)| \tag{2.5}$$

dengan $F_0(y)$ merupakan fungsi distribusi kumulatif yang dihipotesiskan dan $F_n(y)$ merupakan fungsi distribusi kumulatif yang diamati. Untuk kriteria pengujiannya yaitu tolak H_0 jika $D_{hitung} > D_{tabel}(\alpha)$ atau nilai signifikansi $< \alpha$.

Suatu variabel random diskrit y dikatakan berdistribusi Poisson dengan parameter μ jika dan hanya jika fungsi probabilitasnya berbentuk sebagai berikut:

$$P(y|\mu) = \frac{e^{-\mu}\mu^y}{y!}; y = 0,1,2,\ldots \tag{2.6}$$

ean dan variansi untuk distribusi Poisson diasumsikan sama atau dapat yatakan sebagai $E[Y] = Var[Y] = \mu$. Untuk membuktikan distribusi nerupakan anggota keluarga eksponensial agar dapat melibatkan GLM maka perlu dibuktikan Persamaan (2.6) agar mengikuti model pada Persamaan (2.4) sebagai berikut:

$$f(y|\mu) = \frac{e^{-\mu}\mu^{y}}{y!}$$

$$f(y|\mu) = \exp([y \ln \mu - \mu] - \ln y!) \tag{2.7}$$

dengan $y = a(y)$, $\ln \mu = b(\theta)$, $-\mu = c(\theta)$, dan $-\ln y! = d(y)$. Sehingga distribusi Poisson dapat dikatakan bagian dari anggota distribusi keluarga eksponensial karena Persamaan (2.7) mengikuti model pada Persamaan (2.4).

Regresi Poisson menggunakan GLM agar model dapat digunakan dalam observasi saat variabel respon tidak perlu berdistribusi normal (Cahyandari, 2014). Dalam GLM terdapat komponen fungsi g yang menghubungkan rata-rata dari variabel respon dengan sebuah variabel prediktor linier, yaitu:

$$g(\mu_i) = \eta$$

$$g(\mu_i) = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ji} \tag{2.8}$$

dengan g disebut fungsi penghubung, $i = 1,2, \dots, n$ dan $j = 1,2, \dots, m$. Apabila μ_i merupakan rata-rata dalam suatu periode dengan μ_i diasumsikan tidak berubah dari titik data ke titik data yang lain secara bebas, maka μ_i dapat dimodelkan sebagai fungsi dari variabel prediktor (Irwan dkk., 2021). Fungsi penghubung log dapat diaplikasikan pada model regresi Poisson karena fungsi penghubung log memastikan nilai variabel yang diharapkan dari variabel responnya akan bernilai tidak negatif. Berikut Persamaan (2.9) yang menunjukkan fungsi penghubung log pada model regresi Poisson yang jika dinyatakan dalam bentuk matriks adalah sebagai berikut (Famoye dkk., 2004):

$$\eta = \ln(\mu_i) = \boldsymbol{X}_{ji}^{T}\boldsymbol{\beta}$$

$$\mu_i = \exp(\boldsymbol{X}_{ji}^{T}\boldsymbol{\beta}) = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_m x_{mi}) \tag{2.9}$$

sehingga model regresi Poisson dapat ditulis sebagai berikut:

$$\mu_i = \exp\left(\boldsymbol{X}_{ji}^{T}\boldsymbol{\beta}\right) + \varepsilon_i, i = 1,2, \dots, n \; dan \; j = 1,2, \dots, m \tag{2.10}$$

$\boldsymbol{\beta}$ ditaksir dengan $\widehat{\boldsymbol{\beta}}$ sehingga menghasilkan model sebagai berikut:

$$\hat{\mu}_i = \exp\left(\boldsymbol{X}_{ji}^{T}\widehat{\boldsymbol{\beta}}\right) + \varepsilon_i, i = 1,2, \dots, n \; dan \; j = 1,2, \dots, m \tag{2.11}$$

2.4 *Spline*

Spline merupakan potongan (*truncated*) polinomial tersegmen yang kontinu dan memiliki sifat fleksibilitas sehingga mampu mengatasi pola data yang menunjukkan naik atau turun dengan bantuan knot, serta kurva yang dihasilkan relatif mulus (Asriyanti, 2022). Knot adalah titik yang menunjukkan terjadinya perubahan-perubahan perilaku kurva pada interval-interval yang berbeda (Fathurahman, 2011). Bentuk umum fungsi *spline* dengan orde ke-q adalah sebagai berikut:

$$f(x_i) = \sum_{l=0}^{q} \beta_l x_i^l + \sum_{k=1}^{r} \beta_{q+k}(x_i - \xi_k)_+^q \tag{2.12}$$

dengan x_i adalah faktor prediktor, β adalah parameter, l adalah orde dengan $l = 0,1,2,\dots,q$, ξ adalah titik knot dengan $k = 1,2,\dots,r$, dan $(x_i - \xi_k)_+^q$ adalah fungsi *truncated* yang menunjukkan segmentasi pada data (Insiro dkk., 2023).

Dari bentuk matematis *spline* pada Persamaan (2.12) dapat dikatakan bahwa *spline* merupakan model polinomial yang tersegmen (*piecewise polinomial*) walaupun *spline* bersifat kontinu pada knot-knotnya. *Spline* orde ke-q dapat juga diartikan sebagai model polinomial orde ke-q pada setiap segmennya sehingga *spline* merupakan suatu fungsi potongan polinomial yang setiap fungsinya tergabung dengan titik knot yang menjamin sifat kontinuitas (Eubank, 1999).

Apabila diberikan model regresi nonparametrik seperti pada Persamaan (2.1), maka bentuk umum regresi *spline* keluarga polinomial *truncated* orde ke-q akan menjadi persamaan sebagai berikut (Islamiyati, 2017)::

$$y_i = \sum_{l=0}^{q} \beta_l x_i^l + \sum_{k=1}^{r} \beta_{q+k}(x_i - \xi_k)_+^q + \varepsilon_i \tag{2.13}$$

dengan fungsi *truncated* sebagai berikut:

$$(x_i - \xi_k)_+^q = \begin{cases} (x_i - \xi_k) & ; x_i \geq \xi_k \\ 0 & ; x_i < \xi_k \end{cases}$$

2.5 *Penalized Spline*

...lized Spline merupakan salah satu bentuk penaksir dari regresi ...etrik yang hasil estimasinya diperoleh dengan meminimumkan ... *least square* (PLS), yaitu kriteria estimasi model yang menggabungkan

antara *goodness of fit* dengan fungsi penalti (Islamiyati, 2019). Berikut merupakan bentuk dari fungsi PLS:

$$PLS = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2 + \lambda \sum_{k}^{r} \beta_{q+k}^2 \tag{2.14}$$

dengan λ merupakan parameter penghalus prediktor (Ruppert dkk., 2003). Sebelum meminimumkan fungsi PLS, Persamaan (2.14) akan diubah ke dalam bentuk matriks sebagai berikut:

$$\boldsymbol{PLS} = \frac{1}{n}\left(\boldsymbol{Y} - \boldsymbol{X_{ji}}(\xi_{jk})\boldsymbol{\beta_j}\right)^{\boldsymbol{T}}\left(\boldsymbol{Y} - \boldsymbol{X_{ji}}(\xi_{jk})\boldsymbol{\beta_j}\right) + \lambda_j \boldsymbol{\beta_j^T D_j \beta_j}$$

$$\boldsymbol{PLS} = \frac{1}{n}\left(\boldsymbol{Y^T Y} - 2\boldsymbol{\beta_j^T X_{ji}^T}(\xi_{jk})\boldsymbol{Y} + \boldsymbol{\beta_j^T X_{ji}^T}(\xi_{jk})\boldsymbol{X_{ji}}(\xi_{jk})\boldsymbol{\beta_j}\right) + \lambda_j \boldsymbol{\beta_j^T D_j \beta_j} \tag{2.15}$$

dengan $\boldsymbol{D_j}$ didefinisikan sebagai matriks diagonal sebagai berikut:

$$\boldsymbol{D_j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathrm{I} \end{bmatrix}$$

dengan **I** merupakan matriks identitas untuk $\beta_{(q+1)}, \beta_{(q+2)}, \dots, \beta_{(q+k)}$.

Selanjutnya menghitung nilai diferensiasi dari fungsi **PLS** pada Persamaan (2.15) terhadap $\boldsymbol{\beta_j}$ dan hasilnya disamakan dengan 0, sehingga diperoleh $\boldsymbol{\hat{\beta}_j}$ sebagai berikut:

$$\frac{\boldsymbol{\partial PLS}}{\boldsymbol{\partial \beta_j}} = 0$$

$$\frac{1}{n}\left(0 - 2\boldsymbol{X_{ji}^T}(\xi_{jk})\boldsymbol{Y} + \boldsymbol{2X_{ji}^T}(\xi_{jk})\boldsymbol{X_{ji}}(\xi_{jk})\boldsymbol{\beta_j} + 2\lambda_j \boldsymbol{D_j \beta_j}\right) = 0$$

$$\boldsymbol{\hat{\beta}_j} = \left(\boldsymbol{X_{ji}^T}(\xi_{jk})\boldsymbol{X_{ji}}(\xi_{jk}) + n\lambda_j \boldsymbol{D_j}\right)^{-1} \boldsymbol{X_{ji}^T}(\xi_{jk})\boldsymbol{Y} \tag{2.16}$$

Selanjutnya mensubstitusi nilai $\boldsymbol{\hat{\beta}_j}$ pada model sebagai berikut:

$$\boldsymbol{\hat{y}} = \boldsymbol{X_j}(\xi_{jk})\left(\boldsymbol{X_{ji}^T}(\xi_{jk})\boldsymbol{X_{ji}}(\xi_{jk}) + n\lambda_j \boldsymbol{D_j}\right)^{-1} \boldsymbol{X_{ji}^T Y} \tag{2.17}$$

atau dapat juga ditulis dalam persamaan berikut:

$$\boldsymbol{\hat{y}} = \boldsymbol{H}(\lambda_j)\boldsymbol{Y} \tag{2.18}$$

dengan $\boldsymbol{H}(\lambda_j) = \boldsymbol{X_{ji}}(\xi_{jk})\left(\boldsymbol{X_{ji}^T}(\xi_{jk})\boldsymbol{X_{ji}}(\xi_{jk}) + n\lambda_j \boldsymbol{D_j}\right)^{-1} \boldsymbol{X_{ji}^T}$ (Ruppert, 2002).

ritma *Full Search*

apat orde, jumlah knot, dan parameter penghalus dalam model regresi *l Spline* yang dapat ditentukan nilai optimalnya menggunakan algoritma

Full-Search. Orde merujuk pada derajat polinomial yang digunakan pada setiap segmen yang menunjukkan sejauh mana model dapat menyesuaikan dengan pola dalam data. Parameter penghalus (λ) mempunyai pengaruh yang sangat besar dalam model regresi *Penalized Spline* (Ruppert dkk., 2003). Jika λ besar maka estimasi fungsi yang diperoleh akan semakin mulus, sedangkan jika λ kecil maka estimasi fungsi yang diperoleh akan semakin besar atau fungsi-fungsi menjadi semakin fluktuatif.

Jumlah knot merupakan banyaknya titik knot atau titik yang mengalami perubahan perilaku fungsi pada interval yang berlainan (Agustina dkk., 2015). Dalam *Penalized Spline*, titik knot terletak pada sampel kuantil dari nilai *unique* (tunggal) variabel prediktor $\{x_i\}_{i=1}^{n}$. Dengan kata lain, titik knot pada *Penalized Spline* terletak pada nilai-nilai tunggal variabel prediktor $\{x_i\}_{i=1}^{n}$ yang membagi segugus pengamatan menjadi $(k+1)$ bagian yang sama. Oleh karena itu, penentuan jumlah knot sangat berpengaruh dalam menentukan titik knot pada *Penalized Spline* (Ruppert, 2002).

Salah satu metode untuk menentukan model optimal yaitu algoritma *Full Search*. Dalam algoritma *Full Search*, jumlah knot yang akan dihitung tidak dipilih melainkan dihitung semua yaitu $1,2,3,4,5,\ldots$ untuk $k < (n_{uniq} - q - 1)$ dengan n_{uniq} adalah banyaknya nilai *unique* atau nilai kuantil dari variabel prediktor $\{x_i\}_{i=1}^{n}$, kemudian dipilih jumlah knot, orde, serta parameter penghalus optimal yang memiliki nilai GCV (*Generalized Cross Validation*) minimum (Ruppert dkk., 2003).

Menurut Budiantara (2005), GCV merupakan modifikasi dari CV (*Cross Validation*) yang menggunakan metode dalam memilih model berdasarkan pada kemampuan prediksi dari model tersebut. Sehingga fungsi GCV dapat dinyatakan sebagai berikut:

$$GCV(\lambda) = \frac{n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(1 - n^{-1}tr[H(\lambda)])^2} \tag{2.19}$$

dengan $n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = MSE$ atau *Mean Square Error* yang merupakan inerja atas estimator yang sederhana (Tripena, 2011). Adapun langkah- dalam Algoritma *Full-Search* adalah sebagai berikut:

1. Membandingkan nilai $GCV(\lambda)$ untuk $k = 1$ dan $k = 2$ untuk masing-masing parameter penghalus (λ) yang meminimumkan nilai GCV. Apabila nilai $GCV(\lambda)$ untuk $k = 1$ lebih kecil dari nilai $GCV(\lambda)$ untuk $k = 2$, maka algoritma akan berhenti dengan memilih $k = 1$ sebagai jumlah knot. Dan apabila nilai $GCV(\lambda)$ untuk $k = 2$ lebih besar dari $GCV(\lambda)$ untuk $k = 1$, maka algoritma akan dilanjutkan dengan membandingkan nilai $GCV(\lambda)$ untuk $k = 2$ dan $k = 3$.
2. Membandingkan nilai $GCV(\lambda)$ untuk $k = 2$ dan $k = 3$ untuk masing-masing parameter penghalus (λ) yang meminimumkan nilai GCV. Apabila nilai $GCV(\lambda)$ untuk $k = 2$ lebih kecil dari $GCV(\lambda)$ untuk $k = 3$, maka algoritma akan berhenti dengan memilih $k = 2$ sebagai jumlah knot. Dan apabila nilai $GCV(\lambda)$ untuk $k = 3$ lebih besar dari nilai $GCV(\lambda)$ untuk $k = 2$, maka algoritma akan dilanjutkan dengan membandingkan nilai $GCV(\lambda)$ untuk $k = 3$ dan $k = 4$.
3. Membandingkan nilai $GCV(\lambda)$ untuk $k = 3$ dan $k = 4$ untuk masing-masing parameter penghalus (λ) yang meminimumkan nilai GCV, dan seterusnya.
4. Setelah mendapatkan jumlah knot yang optimal, selanjutnya embandingkan nilai $GCV\ (\lambda)$ untuk orde $1,2,3,\cdots,q$ (Ruppert dkk., 2003).

2.7 *Maximum Likelihood Estimation*

Dalam menaksir parameter model Regresi Poisson tidak dapat menggunakan metode kuadrat terkecil karena terdapat beberapa pelanggaran asumsi mengenai galat yang tidak berdistribusi normal dan variansi galat yang tidak homogen maka digunakan metode *Maximum Likelihood Estimation* (MLE) untuk mengatasi masalah tersebut. MLE merupakan pendekatan yang efektif untuk mengestimasi parameter dalam model statistik dengan mencari nilai-nilai parameter yang memaksimalkan fungsi *likelihood* (Ginting, 2023). Fungsi *likelihood* ini menggambarkan probabilitas dari pengamatan data yang sebenarnya dalam model regresi Poisson. MLE memberikan estimasi yang konsisten dan efisien sehingga metode yang sesuai untuk memodelkan hubungan antara variabel dan variabel respon dalam regresi Poisson karena sudah diketahui fungsi ya (Davis dkk., 2005).

Misalkan $y_1, y_2, \dots, y_n$ adalah sampel acak dari populasi dengan fungsi probabilitas $f(y|\boldsymbol{\theta})$ yang bergantung pada $\boldsymbol{\theta} = \mu, \boldsymbol{\beta}$ dengan $\boldsymbol{\theta}$ adalah parameter yang tidak diketahui. Karena $y_1, y_2, \dots, y_n$ saling bebas, maka fungsi peluang bersama dari $y_1, y_2, \dots, y_n$ dapat dinyatakan sebagai berikut:

$$f(y_1, y_2, \dots, y_n\,|\boldsymbol{\theta}) = f(y_1\,|\boldsymbol{\theta}) \cdot f(y_2\,|\boldsymbol{\theta}) \cdot \dots \cdot f(y_n\,|\boldsymbol{\theta}) \tag{2.20}$$

Hogg, (2013) mendefinisikan fungsi *likelihood* sebagai fungsi peluang bersama dari $y_1, y_2, \dots, y_n$ yang dianggap fungsi dari $\boldsymbol{\theta}$ yang dituliskan sebagai $L(\boldsymbol{\theta}|y)$, yaitu:

$$L(\boldsymbol{\theta}|y) = f(y_1, y_2, \dots, y_n\,|\boldsymbol{\theta})$$

$$L(\boldsymbol{\theta}|y) = f(y_1|\boldsymbol{\theta}) \cdot f(y_2|\boldsymbol{\theta}) \cdot \dots \cdot f(y_n\,|\boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_i|\boldsymbol{\theta}) \tag{2.21}$$

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

$$l(\boldsymbol{\theta}|y) = \ln L(\boldsymbol{\theta}|y)$$

$$l(\boldsymbol{\theta}|y) = \ln\left[\prod_{i=1}^{n} f(y_i\,|\boldsymbol{\theta})\right] = \sum_{i=1}^{n} \ln f(y_i\,|\boldsymbol{\theta}) \tag{2.22}$$

Untuk memperoleh nilai $\boldsymbol{\theta}$ yang memaksimumkan fungsi $l(\boldsymbol{\theta}|y)$, maka $l(\boldsymbol{\theta}|y)$ diturunkan terhadap $\boldsymbol{\theta}$ dan kemudian menyamakannya dengan nol sebagai berikut:

$$l'(\boldsymbol{\theta}|y) = \frac{\partial l(\boldsymbol{\theta}|y)}{\partial \boldsymbol{\theta}} = 0 \tag{2.23}$$

2.8 Algoritma *Fisher Scoring*

Algoritma *Fisher Scoring* adalah salah satu bentuk pengembangan dari metode *Newton Raphson* dengan mengganti matriks *Hessian* dengan matriks informasi (Schworer & Hovey, 2004). Berikut persamaan algoritma *Fisher Scoring*:

$$\hat{\boldsymbol{\theta}}^{(t+1)} = \hat{\boldsymbol{\theta}}^{(t)} + I\left(\boldsymbol{\theta}^{(t)}\right)^{-1} l'\left(\boldsymbol{\theta}^{(t)}\right) \tag{2.24}$$

dengan $I^{(t)}$ adalah taksiran ke-t dari matriks informasi yang diamati. Matriks informasi adalah negatif dari nilai ekspektasi matriks diferensiasi kedua fungsi logaritma natural *likelihood* yaitu $I = -E\left(\frac{\partial^2 l(\boldsymbol{\theta}|y)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}}\right)$.

Adapun langkah-langkah dalam algoritma *Fisher Scoring* adalah sebagai berikut:

1. Menentukan nilai awal $\widehat{\boldsymbol{\theta}}^{(0)}$.
2. Menentukan $l'(\boldsymbol{\theta}^{(t)})$ dan $I(\boldsymbol{\theta})^{(t)}$.
3. Menghitung estimator parameter untuk $t = 0,1,2,\ldots$ dengan menggunakan Persamaan (2.24).
4. Mengulangi iterasi hingga diperoleh nilai yang konvergen, yaitu $\left|\boldsymbol{\theta}^{(t+1)} - \widehat{\boldsymbol{\theta}}^{(t)}\right| \leq \epsilon$ dengan ϵ yaitu bilangan positif yang sangat kecil (Y. Wang, 2007).

2.9 Pengujian Signifikansi Parameter Model Regresi

Terdapat uji dalam mengetahui signifikansi parameter model regresi, yaitu:

1. Uji Simultan

Uji simultan yang digunakan yaitu Uji *Likelihood Ratio* dilakukan untuk menguji signifikansi pengaruh dari semua variabel prediktor secara keseluruhan terhadap variabel respon.

Hipotesis:

H_0: $\beta_{11} = \beta_{12} = \cdots = \beta_{j(q+k)} = 0$

H_1: Minimal ada satu $\beta_{j(q+k)} \neq 0,\ j = 1,2,\ldots,m,\ l = 1,2,\ldots,q,\ k = 1,2,\ldots,r$

Statistik uji:

$$G = -2ln\left[\frac{L(\widehat{\boldsymbol{\theta}}_0)}{L(\widehat{\boldsymbol{\theta}})}\right] = -2\left[L(\widehat{\boldsymbol{\theta}}_0) - L(\widehat{\boldsymbol{\theta}})\right] \tag{2.25}$$

dengan $L(\widehat{\boldsymbol{\theta}}_0)$ merupakan log *likelihood* pada model yang tidak mengandung variabel respon dan $(\widehat{\boldsymbol{\theta}})$ merupakan log *likelihood* pada model yang mengandung seluruh variabel respon (Sundari & Sihombing, 2021).

2. Uji Parsial

rsial yang digunakan yaitu Uji *Wald* yang merupakan suatu pengujian ujuan mengetahui masing-masing koefisien regresi variabel prediktor signifikan atau tidak terhadap variabel respon. Dalam uji ini digunakan hipotesis sebagai berikut,

Hipotesis:

H_0: $\beta_{j(q+k)} = 0,\ j = 1, 2, \ldots, m,\ l = 1, 2, \ldots, q,\ k = 1, 2, \ldots, r$

H_1: $\beta_{j(q+k)} \neq 0,\ j = 1, 2, \ldots, m,\ l = 1, 2, \ldots, q,\ k = 1, 2, \ldots, r$

Statistik uji:

$$W_{j(q+k)} = \left[\frac{\hat{\beta}_{j(q+k)}}{SE\left(\hat{\beta}_{j(q+k)}\right)}\right]^2 \tag{2.26}$$

dengan $\hat{\beta}_{j(q+k)}$ adalah nilai estimasi parameter dan $SE(\hat{\beta}_{j(q+k)})$ adalah standar *error* dari nilai estimasi parameter (Dewanti dkk., 2016).

2.10 Jumlah Kematian Neonatus

Neonatus atau bayi baru lahir yang berusia 0 hingga 28 hari merupakan kelompok umur yang sangat rentan untuk kelangsungan hidup bayi karena mudah terkena penyakit (Sembiring, 2019). Kematian neonatus menjadi semakin krusial karena proporsi kematian neonatus meningkat di seluruh dunia selama 25 tahun terakhir serta mendominasi dari jumlah kematian anak dibawah usia lima tahun. Selain itu, intervensi kesehatan yang dibutuhkan untuk mengatasi penyebab utama kematian neonatus berbeda dari yang diperlukan untuk mengatasi kematian anak dibawah usia lima tahun. Oleh karena itu, penting adanya kesadaran terhadap kesehatan ibu hamil dan bayi saat lahir (Dinas Kesehatan Sulawesi Selatan, 2021).