PEMODELAN *GENERALIZED POISSON REGRESSION* DENGAN ESTIMASI *MODIFIED JACKKNIFED POISSON RIDGE* PADA KASUS JUMLAH KEMATIAN IBU DI PROVINSI SULAWESI SELATAN TAHUN 2021

NURUL HIDAYA NASIR

H051201042

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

2024

PEMODELAN *GENERALIZED POISSON REGRESSION* DENGAN ESTIMASI *MODIFIED JACKKNIFED POISSON RIDGE* PADA KASUS JUMLAH KEMATIAN IBU DI PROVINSI SULAWESI SELATAN TAHUN 2021

NURUL HIDAYA NASIR
H051201042

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

PEMODELAN *GENERALIZED POISSON REGRESSION* DENGAN ESTIMASI *MODIFIED JACKKNIFED POISSON RIDGE* PADA KASUS JUMLAH KEMATIAN IBU DI PROVINSI SULAWESI SELATAN TAHUN 2021

NURUL HIDAYA NASIR
H051201042

Skripsi

Diajukan sebagai salah satu syarat untuk mencapai gelar Sarjana Statistika

Program Studi Statistika

pada

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

SKRIPSI

PEMODELAN *GENERALIZED POISSON REGRESSION* DENGAN ESTIMASI *MODIFIED JACKKNIFED POISSON RIDGE* PADA KASUS JUMLAH KEMATIAN IBU DI PROVINSI SULAWESI SELATAN TAHUN 2021

NURUL HIDAYA NASIR
H051201042

Skripsi,

telah dipertahankan di hadapan Panitia Ujian Sarjana Statistika pada 2 Agustus 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

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

Mengesahkan:	Mengetahui,
Pembimbing tugas akhir,	Ketua Program Studi,
Sitti Sahriman, S.Si., M.Si.	Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19881018 201504 2 002	NIP. 19770808 200501 2 002

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN UNIVERSITAS HASANUDDIN

PERNYATAAN KEASLIAN SKRIPSI
DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Generalized Poisson Regression* dengan Estimasi *Modified Jackknifed Poisson Ridge* pada Kasus Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Tahun 2021" adalah benar karya saya dengan arahan dari pembimbing skripsi saya (Ibu Sitti Sahriman, S.Si., M.Si.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. 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 skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 2 Agustus 2024

METERAI TEMPEL
E9787AKX174997955

Nurul Hidaya Nasir
NIM H051201042

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam*, yang telah membawa kita dari zaman kegelapan menuju zaman yang terang benderang. *Alhamdulillaahirobbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh Allah *Subhanahu Wa Ta'ala*, penelitian ini dapat terlaksana dan terselesaikan atas bimbingan, diskusi dan arahan dari Ibu **Sitti Sahriman, S.Si., M.Si.** selaku pembimbing sekaligus penasehat akademik yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya. Ucapan terima kasih juga penulis sampaikan kepada Ibu **Anisa, S.Si., M.Si.** dan Ibu **Dr. Anna Islamiyati, S.Si., M.Si.** atas saran dan kritikan dalam penyempurnaan skripsi ini serta waktu yang telah diberikan kepada penulis. Terima kasih juga kepada pimpinan Universitas Hasanuddin, Ketua Departemen Statistika, para dosen dan staff yang telah memberikan ilmu dan fasilitas kepada penulis.

Ucapan terima kasih serta penghargaan setinggi-tingginya penulis haturkan kepada orang tua tercinta, Ayahanda **Muh. Nasir** dan Ibunda **Hasnah Huseng** yang selalu berjuang dalam mengupayakan yang terbaik untuk penulis dan telah memberikan dukungan, pengorbanan, kasih sayang serta doa yang tak henti-hentinya dipanjatkan kepada penulis. Terima kasih juga kepada kakak-kakak tersayang penulis **Anggi Paramita** dan **Nur Fadilah Nasir** yang selalu memotivasi, memberikan semangat, dukungan, dan doa mulia kepada penulis.

Penulis juga berterima kasih kepada teman-teman **Statistika 2020** terkhusus **Radia Sultan**, **Nurhaliza Rais**, **Muh. Ridwan**, **Fadlan Amin**, **A. Muh. Faldi Fadhil A. Darsyad**, dan **Muhammad Edward**, teman-teman seperbimbingan **Ruslinda**, **Najlah Fauziah, Alisha Shafa Azzahra, Nur Aviatul Zahra, dan Nurfajriyani** yang menjadi tempat diskusi dan senantiasa memberikan semangat, teman-teman **Pois20n** dan warga **Exsite20** yang juga selalu memberikan dukungan. **Nanda Zahra Afifah** terima kasih telah menjadi tempat berkeluh kesah dan senantiasa menemani serta menyemangati penulis serta semua pihak yang telah banyak membantu yang tidak dapat disebutkan satu-persatu, semoga segala dukungan dan partisipasi yang diberikan bernilai ibadah disisi Allah *Subhanahu Wa Ta'ala*. Terakhir, terima kasih untuk diri sendiri yang telah berjuang dan bertahan untuk menyelesaikan skripsi ini, terima kasih sudah menjadi kuat dan tetap waras.

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 berbagi pihak.

Makassar, 2 Agustus 2024

Nurul Hidaya Nasir

ABSTRAK

NURUL HIDAYA NASIR. **Pemodelan *Generalized Poisson Regression* dengan Estimasi *Modified Jackknifed Poisson Ridge* pada Kasus Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Tahun 2021** (dibimbing oleh ibu Sitti Sahriman, S.Si., M.Si.)

Latar Belakang. *Generalized Poisson Regression* (GPR) merupakan pengembangan dari regresi Poisson yang digunakan untuk menganalisis data yang mengalami overdispersi yaitu nilai variansi lebih besar dari rata-rata yang tidak dapat ditangani oleh regresi Poisson. Estimasi parameter GPR umumnya menggunakan *Maximum Likelihood Estimation* (MLE), namun metode ini juga mengasumsikan tidak adanya multikolinearitas pada data yang dapat menyebabkan estimasi parameter menjadi tidak akurat. Untuk menyelesaikan masalah tersebut, digunakan estimasi *Modified Jackknifed Poisson Ridge* (MJPR). Metode ini menggunakan teknik *Jackknife* untuk mengurangi bias dan menangani multikolinearitas, sehingga menghasilkan estimasi parameter yang lebih akurat dan stabil. **Tujuan.** Penelitian ini bertujuan memperoleh faktor-faktor yang mempengaruhi jumlah kematian ibu di Provinsi Sulawesi Selatan tahun 2021 menggunakan model GPR dengan estimasi MJPR. **Metode.** Analisis data dilakukan dalam dua tahap, pertama mengestimasi parameter GPR dengan MLE untuk menangani overdispersi. Kedua mengestimasi parameter GPR dengan MJPR untuk mengatasi multikolinearitas menggunakan beta dari hasil MLE sebagai penduga awal dalam MJPR. **Hasil.** Hasil analisis menunjukkan bahwa model GPR dengan MJPR lebih baik dalam menangani overdispersi dan multikolinearitas berdasarkan nilai AIC (133.1317) yang lebih kecil dibandingkan model GPR dengan MLE (134.9758). Selain itu, estimasi MJPR mampu menjelaskan keragaman data lebih baik berdasarkan nilai R^2 (89.05%) dibandingkan dengan MLE yang memiliki nilai R^2 (88.99%). **Kesimpulan.** Faktor-faktor yang teridentifikasi mempengaruhi jumlah kematian ibu di Provinsi Sulawesi Selatan berdasarkan hasil pengujian signifikansi parameter yaitu variabel Perkiraan Jumlah Ibu Hamil dengan Komplikasi Kebidanan (X_4) pada taraf signifikan 0.05.

Kata Kunci: *Generalized Poisson Regression, Maximum Likelihood Estimation, Modified Jackknifed Poisson Ridge, Multikolinearitas, Overdispersi.*

ABSTRACT

NURUL HIDAYA NASIR. **Modeling Generalized Poisson Regression with Modified Jackknifed Poisson Ridge Estimation for Maternal Deaths in South Sulawesi Province in 2021** (supervised by Mrs. Sitti Sahriman, S.Si., M.Si.)

Background. Generalized Poisson Regression (GPR) is a development of Poisson regression which is used to analyze data that experiences overdispersion, where the variance value is greater than the mean, which cannot be handled by Poisson regression. GPR parameter estimation generally uses Maximum Likelihood Estimation (MLE), but this method also assumes the absence of multicollinearity in the data, which can cause parameter estimates to be inaccurate. To solve this problem, Modified Jackknifed Poisson Ridge (MJPR) estimation is used. This method uses Jackknife technique to reduce bias and handle multicollinearity, resulting in more accurate and stable parameter estimates. **Objective**. This study aimed to identify the factors influencing the number of maternal mortality in South Sulawesi Province in 2021 using the Generalized Poisson Regression (GPR) model with Modified Jackknifed Poisson Ridge (MJPR) estimation. **Methods**. The data analysis was conducted in two stages. First, the parameters of the GPR were estimated using Maximum Likelihood Estimation (MLE) to address overdispersion. Second, the parameters of the GPR were estimated using MJPR to address multicollinearity, used beta from the MLE results as an initial estimates in MJPR. **Results**. The analysis results showed that the GPR model with MJPR was better handling overdispersion and multicollinearity based on a lower AIC value (120.9704) compared to the GPR model with MLE (122.6907). Additionally, the estimation with MJPR was able to explain was better data diversity based on the R^2 value (75.56%) compared to MLE which has an R^2 value (75.28%). **Conclusion.** The factors identified as affecting the number of maternal deaths in South Sulawesi Province, based on the significance testing of parameters, were the variable Estimate of Pregnant Women with Obstetric Complications (X_4) at a significance level of 0.05.

Keywords: Generalized Poisson Regression, Maximum Likelihood Estimation, Modified Jackknifed Poisson Ridge, Multicollinearity, Overdispersion.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Equidispersi	Kesamaan antara nilai rata-rata dan variansi
Overdispersi	Nilai variansi lebih besar dari rataan
Underdispersi	Nilai variansi lebih kecil dari rataan
Fisher scoring	Metode numerik yang digunakan untuk menemukan estimasi parameter maksimum likelihood (MLE) dalam analisis statistik
Vektor gradien	Turunan pertama dari suatu fungsi terhadap vektor parameter
Matriks Hessian	matriks turunan kedua dari suatu fungsi
Matriks informasi fisher	Negatif dari nilai ekspektasi matriks Hessian dari fungsi *loglikelihood yang* digunakan dalam teori estimasi untuk menghitung variansi dari estimator maksimum *likelihood*.
Eksponensial	Fungsi matematika yang dasarnya adalah pangkat eksponen dari bilangan tertentu
Parameter	Konstanta yang digunakan untuk mendefinisikan karakteristik suatu populasi
Multikolinearitas	Kondisi di mana terdapat hubungan (korelasi) antara dua atau lebih variabel prediktor dalam suatu model regresi
Standar error	Ukuran dari seberapa akurat rata-rata sampel mewakili rata-rata populasi.
Ridge	Metode statistik yang digunakan untuk menganalisis data yang mengalami multikolinearitas
Jackknife	Teknik statistik yang digunakan untuk mengestimasi bias dan variansi dari suatu estimator, serta untuk menguji stabilitas model.
Diskrit	Variabel atau data yang dapat mengambil nilai tertentu dan terbatas
Underestimate	Nilai prediksi dari model lebih kecil daripada nilai sebenarnya
Overestimate	Nilai prediksi dari model lebih besar daripada nilai sebenarnya
Centering	Perbedaan antara masing-masing pengamatan dan rata-rata dari semua pengamatan untuk variabel.
Scaling	Gambaran pengamatan kesatuan (unit) standar deviasi dari pengamatan untuk variabel.

DAFTAR LAMBANG/SINGKATAN

Istilah	Arti dan Penjelasan
Y	Variabel respon
X	Variabel prediktor
β	Parameter regresi
α	Parameter dispersi
ε_i	Residual ke-i
η	Vektor prediksi
exp	Exponen
$g(.)$	Fungsi penghubung
n	Jumlah pengamatan
p	Jumlah variabel bebas
c	Jumlah parameter
$F_n(Y)$	Fungsi distribusi kumulatif yang diamati
$F_0(Y)$	Fungsi distribusi kumulatif yang dihipotesiskan
μ	Rata-rata
D^2	Nilai deviance
χ^2	Nilai *Pearson Chi-Squares*.
$H\left(\hat{\theta}^r\right)$	Matriks hessian
$I(\theta^r)$	Matriks informasi fisher
$l'\left(\theta^{(r)}\right)$	Vektor gradien
$\hat{\alpha}^2{}_{max}$	Nilai maksimum dari $\left(\boldsymbol{G}^{\boldsymbol{T}}\widehat{\boldsymbol{\beta}}^*_{ML}\right)^2$
$\hat{\beta}^*_{ML}$	Nilai estimasi *Maksimum Likelihood*
G	*Vector eigen*
Λ_{PR}	Matriks diagonal dari nilai eigen
$\hat{\beta}^*_{MJPR}$	Nilai estimasi *Modified Jackknifed Poisson Ridge*
R_j^2	Koefisien determinasi antara variabel prediktor ke-*j*
$Se(\widehat{\beta}_J)$	Nilai standar *error* parameter regresi ke-*j*
$l\left(\hat{\beta}\right)$	Nilai *loglikelihood* dari model
GPR	*Generalized Poisson Regression*
MJPR	*Modified Jackknifed Poisson Ridge*
MLE	*Maximum Likelihood Estimation*
GLM	*Generalized Linear Model*
VIF	*Variance Inflation Factor*
SDGs	*Sustainable Development Goals*
AIC	*Akaike Information Criteria*
JKR	Jumlah Kuadrat Regresi
JKG	Jumlah Kuadrat Galat
KTR	Kuadrat Tengah Regresi
KTG	Kuadrat Tengah Galat

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGAJUAN ii
HALAMAN PENGESAHAN iii
PERNYATAAN KEASLIAN SKRIPSI v
UCAPAN TERIMA KASIH v
ABSTRAK vi
ABSTRACT vii
DAFTAR ISTILAH viii
DAFTAR LAMBANG/SINGKATAN ix
DAFTAR ISI x
DAFTAR TABEL xii
DAFTAR LAMPIRAN xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Batasan Masalah 3
1.3 Tujuan Penelitian 3
1.4 Manfaat Penelitian 3
1.5 Teori 3
1.5.1 *Generalized Linear Model* 3
1.5.2 Regresi Poisson 4
1.5.3 Equidispersi 5
1.5.4 Multikolinearitas 7
1.5.5 *Generalized Poisson Regression* 8
1.5.6 Algoritma *Fisher Scoring* 8
1.5.7 *Modified Jackknifed Poisson Ridge* 9
1.5.8 Pemilihan Nilai k 12
1.5.9 Uji Kebaikan Model 12
1.5.10 Uji Signifikansi Parameter 12
1.5.11 Angka kematian ibu 13
BAB II METODOLOGI PENELITIAN 15
2.1 Sumber Data 15

2.2 Variabel Penelitian 15
2.3 Tahapan Analisis 15
BAB III HASIL DAN PEMBAHASAN 17
3.1 Estimasi Parameter Model *Generalized Poisson Regression* 17
3.2 Estimasi Parameter *Modified Jackknifed Poisson Ridge* 19
3.3 Pemodelan Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Menggunakan *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge* 19
3.3.1 Statistik Deskriptif 19
3.3.2 Uji Distribusi Poisson 20
3.3.3 Uji Equidisversi 20
3.3.4 Uji Multikolinearitas 21
3.3.5 Pemodelan *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge* 22
3.3.6 Uji Signifikansi Parameter 25
3.3.7 Uji Kebaikan Model 26
BAB IV KESIMPULAN DAN SARAN 27
4.1 Kesimpulan 27
4.2 Saran 27
DAFTAR PUSTAKA 28
LAMPIRAN 31

DAFTAR TABEL

Tabel 1. Pedoman Derajat Hubungan 7
Tabel 2. Variabel Penelitian 15
Tabel 3. Statistik Deskriptif Data Penelitian 19
Tabel 4. Nilai *Pearson Chi-Square* dan *Deviance*..21
Tabel 5. Pengujian Multikolinearitas 21
Tabel 6. Matriks Korelasi 21
Tabel 7. Estimasi Model GPR dengan MLE..22
Tabel 8. Penduga Parameter $(\boldsymbol{\gamma}ML*)$ 23
Tabel 9. Penduga Parameter *Modified Jackknifed Poisson Ridge* 24
Tabel 10. Analisis Variansi GPR dengan estimasi MJPR 25
Tabel 11. Uji Parsial GPR dengan estimasi MJPR 26
Tabel 12. Perbandingan Nilai AIC..26

DAFTAR LAMPIRAN

Lampiran 1. Data Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Tahun 2021.32
Lampiran 2. Uji *Kolmogorov-Smirnov* 33
Lampiran 3. Uji Equidispersi 34
Lampiran 4. Uji Multikolinearitas 35
Lampiran 5. Data Hasil Transformasi..36
Lampiran 6. Matriks Z 37
Lampiran 7. Pemilihan Nilai *k* 38
Lampiran 8. Uji Kebaikan Model 39
Lampiran 9. Riwayat Hidup Peneliti 40

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kesehatan merupakan faktor penting dalam pembangunan dan kemajuan ekonomi suatu negara karena hal tersebut berdampak signifikan pada kualitas manusia sebagai sumber daya yang produktif (Azizah, 2022). Hal ini juga menjadi salah satu tujuan dari *Sustainable Development Goals* (SDGs) yaitu menjamin kehidupan yang sehat dan meningkatkan kesejahteraan hidup untuk seluruh masyarakat tanpa memandang usia (Bappenas, 2023). Dalam sektor kesehatan, SDGs menetapkan 38 target yang harus dicapai. Namun, masih terdapat beberapa isu yang belum teratasi, termasuk penurunan angka kematian ibu yang menjadi indikator penting dalam menentukan derajat kesehatan masyarakat terutama ibu. Penurunan Angka kematian Ibu (AKI) menjadi salah satu sorotan SDGs dimana telah menetapkan target untuk mengurangi rasio kematian ibu menjadi kurang dari 70 per 100.000 kelahiran hidup pada tahun 2030 (BPS, 2022).

Kematian ibu adalah kematian selama kehamilan atau dalam kurun waktu 42 hari setelah melahirkan yang disebabkan karena kehamilan atau pengelolaannya (Dinkes Sulsel, 2021). Berdasarkan pencatatan program kesehatan keluarga di Kementerian Kesehatan, jumlah kematian ibu di Indonesia meningkat setiap tahunnya dimana pada tahun 2021 terdapat 7.389 kematian ibu yang mengalami peningkatan dari 4.627 kematian di tahun 2020 (Kemenkes RI, 2022). Provinsi Sulawesi Selatan menempati urutan ke-7 di Indonesia dengan jumlah kematian ibu tertinggi pada tahun 2021, yaitu sebanyak 198 orang yang terdiri dari kematian ibu hamil 69 orang, kematian ibu bersalin 46 orang, dan kematian ibu nifas 83 orang. Angka ini menunjukkan peningkatan yang signifikan dibandingkan tahun 2020 yang tercatat sebanyak 133 orang (Dinkes Sulsel, 2022). Tingginya jumlah kematian ibu dapat dipengaruhi oleh beberapa faktor diantaranya komplikasi kebidanan, proses persalinan yang tidak dilakukan oleh tenaga kesehatan, serta kurangnya pelayanan kesehatan ibu hamil, bersalin maupun nifas (Adnan, dkk. 2020). Salah satu upaya pencegahan terhadap meningkatnya jumlah kematian ibu adalah dengan memodelkan jumlah kematian ibu dengan faktor-faktor yang diduga mempengaruhinya, sehingga dapat diidentifikasi faktor apa saja yang berpengaruh signifikan terhadap jumlah kematian ibu. Memahami faktor-faktor ini merupakan langkah penting dalam menurunkan jumlah kasus kematian ibu (Prastika, dkk. 2021).

Data mengenai jumlah kematian ibu merupakan data cacahan yang berdistribusi Poisson karena merupakan data dengan peluang kejadiannya kecil pada periode waktu tertentu, sehingga metode statistik yang tepat digunakan untuk memodelkan data adalah regresi Poisson. Untuk melakukan analisis regresi Poisson, salah satu asumsi penting yang harus terpenuhi adalah equidispersi, dimana nilai rata-rata dan variansi memiliki kesamaan (Chaniago dan Wulandari, 2022). Namun, seringkali sulit untuk memastikan bahwa kondisi ini selalu terpenuhi karena data diskrit seringkali mengalami overdispersi atau underdispersi. Data jumlah kematian ibu berpotensi mengalami pelanggaran asumsi equidispersi yang dapat menyebabkan kesalahan standar *error* yang lebih tinggi dan estimasi parameter yang kurang tepat, sehingga dalam penanganannya diperlukan sebuah model regresi yang dapat digunakan untuk mengatasi masalah tersebut. Oleh karena itu, dilakukan pengembangan model menggunakan *Generalized Poisson Regression* untuk mengatasi masalah pelanggaran asumsi equidispersi yang terjadi pada regresi Poisson (Islamiyah, 2020).

Estimasi parameter *Generalized Poisson Regression* umumnya menggunakan metode *Maximum Likelihood Estimation* (MLE), namun metode ini seringkali memberikan hasil yang kurang tepat dan salah satu masalahnya adalah terjadinya multikolinearitas. Multikolinearitas merupakan kondisi di mana terdapat hubungan (korelasi) antara dua atau lebih variabel bebas. Adanya multikolinearitas yang tinggi pada variabel bebas dapat membuat model regresi menjadi tidak tepat dan dapat menghasilkan taksiran parameter memiliki nilai variansi menjadi tidak minimum. Terbatasnya jumlah pengamatan dalam data dan adanya berbagai faktor yang mempengaruhi, mengakibatkan potensi munculnya multikolinearitas pada data kematian ibu. Salah satu cara untuk mengatasi masalah ini adalah dengan menggunakan metode regresi *ridge* yang pertama kali diperkenalkan oleh Hoerl dan Kennard (Munawaroh, 2018).

Penelitian sebelumnya telah dilakukan oleh Chaniago dan Wulandari (2022) mengenai perbandingan metode antara *Generalized Poisson Regression* (GPR) dan *Negative Binomial Regression* (NBR) dalam mengatasi overdispersi. Hasil penelitian menunjukkan bahwa metode yang paling baik digunakan dalam memodelkan jumlah kematian bayi untuk mengatasi overdispersi adalah metode GPR karena memiliki kriteria kebaikan model AIC, AICc, BIC, dan BICc yang lebih kecil dibandingkan dengan metode NBR. Sementara itu, penelitian yang dilakukan Munawaroh (2018) yang berjudul "Estimator Baru *Modified Jackknifed* untuk Mengatasi Multikolinearitas pada Regresi Poisson" menunjukkan bahwa metode *Modified Jackknifed Poisson Ridge* lebih baik dalam mengatasi multikolinearitas dibandingkan dengan metode *Poisson Ridge* dan *Jackknifed Poisson Ridge.* Adapun penelitian Azizah (2022) menggunakan Regresi *Poisson Inverse Gaussian* dengan Estimasi Metode *Modified Jackknifed Poisson Ridge Regression* pada Kasus Angka Kematian Ibu di Jawa Timur menyatakan bahwa data angka kematian ibu mengalami overdispersi, sehingga pemodelan regresi *Poisson Inverse Gaussian* dapat digunakan untuk analisis faktornya dan estimasi dengan *Modified Jackknifed Poisson Ridge* dapat digunakan untuk mengatasi multikolinearitas pada kasus tersebut karena penaksiran parameter dengan *Maximum Likelihood* dapat memberikan hasil yang kurang tepat.

Oleh karena itu, berdasarkan uraian di atas penulis terdorong untuk melakukan penelitian menggunakan metode *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge* untuk mengatasi masalah pelanggaran asumsi equidispersi dan masalah multikolinearitas dalam pemodelan faktor-faktor yang mempengaruhi jumlah kasus kematian ibu di Provinsi Sulawesi Selatan. Metode *Modified Jackknifed Poisson Ridge* merupakan metode gabungan antara regresi *Poisson Ridge* dan *Jackknifed Poisson Ridge*.

1.2 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data penelitian yang digunakan yaitu data kasus jumlah kematian ibu di Provinsi Sulawesi Selatan Tahun 2021 yang diperoleh dari Dinas Kesehatan Provinsi Sulawesi Selatan.
2. Algoritma iterasi *fisher scoring* digunakan dalam mengestimasi parameter model *Generalized Poisson Regression.*
3. Uji kebaikan model berdasarkan nilai *Akaike Information Criteria* dan nilai *R-Square* terkecil dari masing-masing model untuk menentukan model terbaik.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini yaitu:

1. Mendapatkan model *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge* pada kasus Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Tahun 2021.
2. Memperoleh faktor-faktor yang secara signifikan mempengaruhi Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Tahun 2021 berdasarkan model *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge*.

1.4 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Sebagai sarana untuk menambah pengetahuan dan wawasan peneliti terkait penerapan metode *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge* dalam mengatasi overdispersi dan multikolinearitas pada data yang diharapkan dapat memberikan referensi bagi penelitian selanjutnya.
2. Sebagai bahan informasi dan pertimbangan bagi masyarakat, tenaga kesehatan dan pemerintah dalam mengambil kebijakan untuk mengatasi masalah angka kematian ibu yang masih tinggi dengan memberikan informasi mengenai faktor-faktor yang mempengaruhi kematian ibu khususnya di Sulawesi Selatan.

1.5 Teori

1.5.1 *Generalized Linear Model*

Generalized Linear Model (GLM) adalah perluasan dari model regresi ketika variabel respon tidak mengharuskan mengikuti distribusi normal tetapi masuk dalam keluarga eksponensial. Menurut McCullagh dan Nelder (1989) dalam GLM terdapat tiga komponen utama, yaitu:

1. Komponen acak, ialah kompenen yang mengidentifikasi distribusi dari variabel respon yang berasal dari keluarga eksponensial. Menurut duson (2002), suatu variabel acak Y dapat dikatakan anggota keluarga distribusi eksponensial jika memiliki fungsi peluang sebagai berikut:

$$f(y) = s(y)t(\theta)e^{a(y)b(\theta)}$$
$$= exp[a(y)b(\theta) + c(\theta) + d(y)] \quad (1)$$

dengan $a(y), b(\theta), s(y) = exp\ d(y), t(\theta) = exp\ c(\theta)$ adalah fungsi yang diketahui dan θ adalah parameter dispersi.

2. Komponen sistematis, ialah komponen yang menentukan variabel-variabel prediktor dari model untuk menjelaskan hubungan antar variabel-variabel dalam sebuah model linear. Dalam notasi matriks ditulis $\eta = X\beta$ dengan η merupakan vektor berukuran $n \times 1$ dari observasi, $\boldsymbol{X}$ adalah matriks berukuran $(n \times p) + 1$ yang elemennya terdiri dari variabel prediktor dan β adalah vektor yang berukuran $(p + 1) \times 1$ yang elemennya terdiri dari parameter regresi,
$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} \quad (2)$$
3. Fungsi penghubung (*link function*), ialah fungsi yang menghubungkan rata-rata dari variabel respon dengan variabel-variabel prediktor melalui persamaan linear. Misalkan $\mu_i = E(Y_i),\ i = 1,2, \dots, n$, fungsi untuk menghubungkan μ_i dengan η_i adalah g sehingga $g(\mu_i) = \eta_i$. Fungsi penghubung g menghubungkan $E(Y_i)$ dengan variabel prediktor, yaitu:
$$g(\mu_i) = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij},\ \ i = 1,2, \dots, n \quad (3)$$

1.5.2 Regresi Poisson

Regresi Poisson adalah suatu metode analisis regresi nonlinier yang dapat digunakan dalam menganalisis data yang variabel responnya bersifat diskrit (*count*) dan diasumsikan berdistribusi Poisson. Distribusi Poisson merujuk pada pola distribusi probabilitas diskrit yang digunakan untuk menggambarkan suatu peristiwa yang jarang terjadi dalam interval waktu atau daerah tertentu (Safira, 2023). Untuk mengetahui suatu data berdistribusi Poisson atau tidak, dapat dilakukan menggunakan uji *Kolmogorov- Smirnov* dengan statistik uji sebagai berikut:
$$D_{hitung} = maks|F_n(Y) - F_0(Y)| \quad (4)$$
dengan:

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

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

Untuk kriteria pengujiannya yaitu tolak H_0 apabila $D_{hitung} > D_{tabel}$ atau nilai signifikansi $< \alpha$.

Menurut (Myers 1990) dalam (Sabtika, dkk. 2021) Jika variabel respon (Y) merupakan data diskrit berdistribusi Poisson dengan parameter $\mu > 0$, maka fungsi probabilitasnya adalah sebagai berikut:
$$f(y, \mu) = \frac{e^{-\mu}\mu^y}{y!}, untuk\ y = 0,1,2, \dots \quad (5)$$
dengan:

y: banyaknya peristiwa yang terjadi selama suatu selang waktu atau daerah tertentu

μ: rata-rata banyaknya peristiwa yang terjadi selama selang waktu atau dalam daerah tertentu

Dalam distribusi Poisson, terdapat situasi dimana rata-rata dan variansinya mempunyai nilai yang sama, yaitu $E[Y] = Var[Y] = \mu$ hal ini dikenal dengan equidispersi. Untuk membuktikan distribusi Poisson merupakan anggota keluarga eksponensial agar dapat melibatkan GLM maka perlu dibuktikan dengan Persamaan (5) agar mengikuti model Persamaan (1) sebagai berikut:

$$\begin{aligned} f(y,\mu) &= \frac{e^{-\mu}\mu^{y}}{y!} \\ &= \exp\left\{ln\left(\frac{e^{-\mu}\mu^{y}}{y!}\right)\right\} \\ &= exp\{ln(e^{-\mu}\mu^{y}) - ln(y!)\} \\ &= exp\{ln(\mu^{y}e^{-\mu}) - ln(y!)\} \\ &= exp\{(y\, ln\, \mu - \mu) - \, ln\, y!\} \end{aligned} \quad (6)$$

dengan $y = a(y), ln\mu = b(\theta), -\mu = c(\theta), -\, ln\, y! = d(y)$. Terbukti bahwa Persamaan (6) mengikuti model pada Persamaan (1), sehingga dapat dikatakan distribusi Poisson merupakan anggota dari distribusi keluarga eksponensial.

Regresi Poisson menggunakan GLM agar model dapat digunakan dalam pengamatan saat variabel respon tidak berdistribusi normal. Dalam GLM terdapat fungsi penghubung (*link function)* $g(.)$ yang menghubungkan nilai rata-rata variabel respon dengan variabel prediktor, yaitu (Sinaga dan Sinulingga, 2021):

$$\begin{aligned} g(\mu_i) &= \eta_i \\ &= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \\ &= \boldsymbol{x}_i^{\boldsymbol{T}}\boldsymbol{\beta} \end{aligned} \quad (7)$$

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

Durmus (2020) menjelaskan bahwa fungsi penghubung antara nilai variabel respon y_i dan variabel prediktor pada model regresi Poisson adalah log, yang digunakan untuk memastikan bahwa y_i selalu memiliki nilai non-negatif. Dengan demikian, model regresi Poisson untuk y_i dapat dirumuskan sebagai berikut (Putri, 2023):

$$\begin{aligned} g(\mu_i) = \ln \mu_i &= \boldsymbol{x}_i^{\boldsymbol{T}}\boldsymbol{\beta} \\ \ln \mu_i &= \boldsymbol{x}_i^{\boldsymbol{T}}\boldsymbol{\beta} \\ \mu_i &= exp\left(\boldsymbol{x}_i^{\boldsymbol{T}}\boldsymbol{\beta}\right) \end{aligned} \quad (8)$$

dengan:

$\boldsymbol{x}_i^{\boldsymbol{T}}$: $[1 \quad X_{i1} \quad X_{i2} \quad \dots \quad X_{ip}]$ menunjukkan vektor yang berukuran $1 \times (p+1)$ dari variabel prediktor yang digunakan untuk memodelkan μ_i

$\boldsymbol{\beta}$: $[\beta_0 \quad \beta_1 \quad \beta_2 \quad \dots \quad \beta_p]^T$ menunjukkan vektor parameter regresi Poisson yang berukuran $(p+1) \times 1$

μ_i : rata-rata varibel prediktor pada lokasi pengamatan ke-*i* yang merupakan ekspektasi dari $Y \sim poisson\ (\mu)$

1.5.3 Equidispersi

Regresi Poisson memiliki asumsi penting yang harus terpenuhi yaitu equidispersi, di mana nilai rata-rata dan variansi dari variabel Y memiliki nilai yang sama. Salah satu penyimpangan umum dari asumsi ini adalah overdispersi, yang terjadi ketika nilai variansi lebih besar nilai rata-rata pada variabel Y (Ismanto, 2019). Overdispersi umumnya muncul karena ketergantungan antar pengamatan atau adanya kecenderungan pada data, yaitu adanya kejadian awal mempengaruhi kejadian berikutnya, sedangkan salah satu asumsi distribusi Poisson adalah bahwa pengamatan tidak tergantung satu sama lain (Moksony dan Hedegus, 2014).

McCullagh dan Nelder (1989) menyatakan bahwa masalah overdispersi akan sering dijumpai dalam analisis data diskrit. Munculnya masalah overdispersi dalam pengamatan data diskrit dapat disebabkan oleh adanya keragaman dalam peluang respon dan adanya korelasi antar variabel respon yang ditunjukkan oleh ketidakbebasan antar variabel respon. Kedua kejadian tersebut merupakan kejadian yang saling berhubungan, artinya jika terdapat keragaman dalam peluang respon, maka terdapat korelasi antar variabel respon. Begitu juga sebaliknya, jika terdapat korelasi antara variabel respon, maka terdapat keragaman dalam peluang respon (Silma dan Hajarisman, 2022). McCullagh dan Nelder (1989) menyatakan bahwa kedua kejadian tersebut dapat terjadi karena adanya pengelompokan (*clustering*) dalam populasi.

Penggunaan metode statistika yang mengasumsikan ketidakbebasan antar variabel respon dapat menjadi tidak tepat, terutama jika terdapat korelasi antar pengamatan. Overdispersi pada regresi Poisson menghasilkan simpangan baku dari estimasi parameter jauh lebih kecil daripada nilai sebenarnya (*underestimate*), sementara signifikansi dari variabel prediktor jauh lebih besar daripada nilai sebenarnya (*overestimate*), sehingga kesimpulan yang dihasilkan menjadi tidak valid (Ismail & Jemain, 2007).

Kondisi tidak terpenuhinya asumsi equidispersi dapat diketahui dengan melihat dari nilai *deviance* atau *person chi-square* (McCullagh dan Nelder, 1989):

1. *Deviance*

$$\phi_1 = \frac{D^2}{db} \tag{9}$$

dengan:

$$D^2 = 2\sum_{i=1}^{n}\left\{y_i \ln\left(\frac{y_i}{\mu_i}\right) - (y_i - \mu_i)\right\}$$

2. *Pearson Chi-Square*

$$\phi_2 = \frac{\chi^2}{db} \tag{10}$$

dengan:

$$\chi^2 = \sum_{i=1}^{n}\frac{(y_i - \mu_i)^2}{var(\mu_i)}$$

Dalam konteks ini, db yaitu derajat bebas $n - c$ dengan n adalah jumlah pengamatan, $c = p + 1$ merupakan jumlah parameter, D^2 merupakan nilai *deviance* dan χ^2 merupakan nilai *pearson chi-square*. Jika nilai ϕ lebih dari 1 maka menunjukkan nilai variansi lebih besar dari rataan sehingga data dikatakan mengalami overdispersi. Namun, jika nilai ϕ kurang dari 1 maka menunjukkan nilai variansi lebih kecil dari rataan sehingga data dikataakan mengalami underdispersi.

1.5.4 Multikolinearitas

Multikolinieritas pertama kali diperkenalkan oleh Ragner Frisch, yaitu adanya hubungan linier yang sangat kuat antara variabel prediktor dalam suatu model regresi (Munawaroh, 2018). Keberadaan multikolinieritas dapat mengakibatkan ketidakstabilan koefisien regresi, sehingga tidak lagi mencerminkan secara akurat pengaruh dari variabel independen. Hal ini juga dapat meningkatkan besarnya nilai *error* dalam estimasi, yang dapat mengganggu validitas analisis (Montgomery, 1992). Pendekatan kuadrat terkecil menjadi kurang tepat karena adanya multikolinieritas. Untuk mengidentifikasi multikolinearitas dapat dilihat dengan menghitung nilai *Variance Inflation Factor* dan melihat korelasi antar variabel prediktornya.

1. *Variance Inflation Factor*

   Variance Inflation Factor (VIF) merupakan salah satu cara untuk mengidentifikasi multikolinieritas pada data. Kolinieritas dianggap tinggi jika VIF memiliki nilai lebih dari 10. Nilai VIF dapat didefinisikan sebagai berikut (Gujarati, 2004):

$$VIF = \frac{1}{\left(1-R_j^2\right)} \tag{11}$$

   dimana R_j^2 merupakan koefisien determinasi antara variabel prediktor ke-*j* dengan variabel prediktor yang lainnya dimana $j = 1,2, \dots, p$.

2. Nilai korelasi

   Korelasi (r) bertujuan untuk mengukuran seberapa kuat hubungan antara dua variabel. Jika koefisien korelasi cukup tinggi maka dapat diduga ada multikolinieritas dalam model. Nilai korelasi dapat dilambangkan dengan persamaan berikut (Sarwono, 2006):

$$\rho = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1p} & r_{2p} & \cdots & 1 \end{bmatrix}$$

$$r_{ij} = \frac{1}{n-2} \sum_{r=1}^{n} \left(\frac{x_{ir}-\bar{X}_i}{\sqrt{S_{ii}}}\right)\left(\frac{x_{jr}-\bar{X}_j}{\sqrt{S_{ii}}}\right) \tag{12}$$

   Keterangan:
   r_{ij} : Korelasi antara variabel X_i dan X_j
   $\bar{X}_i$: Rata-rata variabel X_i
   $\bar{X}_j$: Rata-rata variabel X_j
   S_{ii} : Standar deviasi variabel X_i
   S_{jj} : Standar deviasi variabel X_j
   Untuk i = j menghasilkan r = 1
   Pedoman derajat hubungan antar dua variabel dijelaskan pada Tabel 1 berikut:

Tabel 1. Pedoman Derajat Hubungan

Interval Koefisien	Tingkat Hubungan
0.00-0.19	Sangat Lemah
0.20-0.39	Lemah
0.40-0.59	Sedang
0.60-0.79	Kuat
0.80-1.00	Sangat Kuat

1.5.5 *Generalized Poisson Regression*

Regresi Poisson tidak tepat digunakan dalam memodelkan data yang mengalami overdispersi atau underdispersi. Oleh karena itu, perlu dilakukan pendekatan dengan model regresi yang lebih tepat. *Generalized Poisson Regression* (GPR) adalah suatu model regresi yang digunakan untuk menganalisis hubungan antara variabel respon (Y) yang berupa data diskrit dengan satu atau lebih variabel prediktor (X). GPR merupakan analisis statistik yang dapat digunakan untuk mengatasi masalah equidispersi, underdispersi dan overdispersi pada data cacahan (Islamiyah, 2020). Model GPR memiliki parameter μ dan parameter tambahan dispersi yakni α. *GPR* merupakan pengembangan dari regresi Poisson, di mana variabel respon (Y) memiliki distribusi *generalized Poisson* dengan fungsi distribusi sebagai berikut (Wang dan Famoye, 1997) dalam (Chaniago dan Wulandari, 2022):

$$f(y_i; \mu_i, \alpha) = \left(\frac{\mu_i}{1 + \alpha\mu_i}\right)^{y_i} \frac{(1 + \alpha y_i)^{y_i - 1}}{y_i!} exp\left[\frac{-\mu_i(1 + \alpha y_i)}{1 + \alpha\mu_i}\right] \tag{13}$$

dengan mean dan variansi model GPR adalah $E(y_i) = \mu_i$ dan $Var(y_i) = \mu_i(1 + \alpha\mu_i)^2$ dengan memuat parameter dispersi α. Jika $\alpha = 0$ maka Persamaan (13) akan menjadi fungsi distribusi Poisson. Hal ini menunjukkan bahwa distribusi *generalized Poisson* merupakan perluasan dari distribusi Poisson. Model GPR dapat dinyatakan sebagai berikut (ismail dan jemain, 2007):

$$y_i = \mu_i + \varepsilon_i, i = 1,2, \dots, n \tag{14}$$

dengan $\mu_i = exp\left(\boldsymbol{x}_i^{\boldsymbol{T}}\boldsymbol{\beta}\right)$

1.5.6 Algoritma *Fisher Scoring*

Algoritma *Fisher Scoring* merupakan pengembangan dari algoritma *Newton Raphson* yang digunakan dalam statistik untuk menyelesaikan persamaan *Maximumn Likelihood* dengan mengganti matriks Hessian dengan matriks Informasi. Matriks informasi adalah negatif dari nilai ekspektasi dari matriks turunan kedua fungsi *loglikelihood*. Persamaan iterasi *fisher scoring* adalah sebagai berikut (Wahyuni, 2021):

$$\widehat{\boldsymbol{\theta}}^{(r+1)} = \widehat{\boldsymbol{\theta}}^{(r)} + I\left(\boldsymbol{\theta}^{(r)}\right)^{-1} l'\left(\boldsymbol{\theta}^{(r)}\right) \tag{15}$$

Dengan $I^{(r)}$ adalah taksiran ke-r dari matriks informasi yang diamati. Matriks informasi dalam penulisam ini, yaitu $I\boldsymbol{\theta}^{(r)} = -E\left(\left[H\left(\widehat{\boldsymbol{\theta}}^{r}\right)\right]\right)$ dan $l'^{(r)}$ adalah turunan pertama dari fungsi *loglikelihood.*

Pengestimasian parameter menggunakan algoritma *fisher scoring* dilakukan dengan langkah-langkah sebagai berikut:

1. Menentukan vektor awal parameter $\widehat{\boldsymbol{\theta}}_{(0)}$ menggunakan model regresi Poisson
2. Membentuk vektor gradien $l'\left(\boldsymbol{\theta}^{(r)}\right)$, matriks Hessian $H\left(\widehat{\boldsymbol{\theta}}^{r}\right)$ dan matriks informasi fisher $I(\boldsymbol{\theta}^{r})$
3. Memasukkan nilai $\widehat{\boldsymbol{\theta}}_{(0)}$ sehingga diperoleh vektor gradien $l'\left(\boldsymbol{\theta}^{(r)}\right)$ matriks Hessian $H\left(\widehat{\boldsymbol{\theta}}^{r}\right)$
4. Menghitung estimator parameter untuk r = 0,1,2, … dengan menggunakan Persamaan (15)

5. Mengulangi iterasi hingga diperoleh nilai yang konvergen, yaitu $|\boldsymbol{\theta}^{r+1} - \widehat{\boldsymbol{\theta}}^{r}| \leq \epsilon$ dengan ϵ adalah konstanta positif yang ditentukan.

1.5.7 *Modified Jackknifed Poisson Ridge*

Regresi *Ridge* adalah metode statistik yang pertama kali diperkenalkan oleh Hoerl dan Kennard pada tahun 1970 sebagai solusi untuk mengatasi masalah multikolinearitas dalam model regresi (Winda, 2021). Multikolinearitas terjadi ketika terdapat korelasi tinggi antara beberapa variabel independen dalam model, yang mengakibatkan perkiraan parameter menjadi tidak stabil. Pendekatan regresi *ridge* memodifikasi metode kuadrat terkecil dengan menambahkan tetapan bias kecil (k) pada diagonal utama matriks korelasi variabel independen $\boldsymbol{X^T X}$, sehingga menghasilkan perkiraan parameter yang lebih stabil (Munawaroh, 2018). Kemudian pada tahun 1977, Hinkley mengusulkan estimator *ridge* yang didasarkan pada teknik *jackknife* untuk mengurangi bias pada estimator regresi *ridge*, namun penelitian Gruber (1998) menunjukkan bahwa metode *jackknife* menghasilkan variansi dan MSE yang lebih tinggi dalam beberapa kondisi. Sehingga Batah dan Gore (2008) menyarankan penggunaan metode *modified jackknifed ridge*, yang menghasilkan variansi yang lebih kecil, untuk menangani masalah multikolinearitas. Metode ini menggabungkan metode *ridge* dan *jackknifed ridge* yang kemudian diadaptasi ke regresi Poisson oleh K. Månsson dan G. Shukur (2011).

Pada tahap awal, data yang akan dianalisis dengan regresi *ridge* harus menjalani proses pemusatan dan penskalaan. Pemusatan atau *centering* adalah proses mengurangi setiap nilai dalam suatu variabel dengan nilai rata-ratanya. Tujuan dari pemusatan adalah untuk menghilangkan perbedaan antara setiap pengamatan dengan rata-rata dari semua pengamatan dalam variabel tersebut, sehingga menghasilakan variabel baru dengan nilai rata-rata dari variabel tersebut menjadi nol. Penskalaan atau *scaling* adalah proses mengubah nilai variabel menjadi satuan standar deviasi. Tujuan dari penskalaan adalah untuk menggambarkan pengamatan dalam satuan standar deviasi, sehingga variabel memiliki rentang nilai yang serupa dan lebih mudah untuk dibandingkan (Winda, 2021). Transformasi data dengan metode *centering* dan *scaling* dilakukan untuk mempermudah analisis data yang memiliki satuan berbeda-beda. Namun, pada regresi Poisson, variabel responsnya berupa data diskrit non-negatif, sehingga metode *centering* and *scaling* hanya dilakukan pada variabel prediktor saja. Proses standarisasi dilakukan dengan cara mentransformasikan variabel prediktor X dalam bentuk berikut (Munawaroh, 2018):

$$X^*ij = \frac{Xij - \bar{X}j}{\sqrt{n-1}S_j} \text{ dengan } S_j = \sqrt{\frac{\sum_{i=1}^{n}(Xij - \bar{X}j)^2}{n-1}} \tag{16}$$

dinotasikan X^* debagai matriks X yang telah distandarisasi.

Metode *Modified Jackknifed Poisson Ridge* (MJPR) adalah pengembangan lebih lanjut dari regresi Poisson *ridge* yang menggunakan teknik *jackknife* untuk mengurangi bias pada estimator regresi *ridge* dan untuk menghitung variansi parameter secara lebih stabil. Metode MJPR merupakan metode gabungan antara regresi *Poisson Ridge* (PR) dan regresi *Jackknifed Poisson Ridge* (JPR).

Untuk memperoleh estimator MJPR diperoleh dengan dengan mengikuti penelitian sebelumnya dari Singh, dkk. (Munawaroh, 2018) menggunakan matriks $\boldsymbol{G}$ berukuran $k \times k$ dengan elemen vektor eigen dari $\boldsymbol{X^T\widehat{W}X}$ dan Λ_{PR} yang merupakan matriks diagonal $(\lambda_{1PR}, \dots, \lambda_{kPR})$ dengan λ_{iPR} merupakan nilai eigen dari matriks $\boldsymbol{X^T\widehat{W}X}$ yang bersesuaian matriks $\boldsymbol{G}$, sehingga:

$$\begin{aligned}\Lambda_{PR} &= \boldsymbol{G^T X^T \widehat{W} X G}\\ &= (\boldsymbol{XG})^{\boldsymbol{T}}\boldsymbol{\widehat{W}}(\boldsymbol{XG})\\ &= \boldsymbol{Z^T\widehat{W}Z}\end{aligned} \tag{17}$$

dengan $\boldsymbol{Z} = \boldsymbol{XG}$ dan $\boldsymbol{\hat{s}} = \boldsymbol{X\widehat{\beta}}_{ML}$, maka bentuk $\boldsymbol{\widehat{\gamma}}_{ML}$ dapat ditulis sebagai berikut:

$$\begin{aligned}\boldsymbol{\widehat{\gamma}}_{ML} &= \left[\boldsymbol{Z^T\widehat{W}Z}\right]^{-1}\boldsymbol{Z^T\widehat{W}\hat{s}}\\ &= \left[\boldsymbol{Z^T\widehat{W}Z}\right]^{-1}\boldsymbol{Z^T\widehat{W}X\widehat{\beta}}_{ML}\\ &= \left[\boldsymbol{Z^T\widehat{W}Z}\right]^{-1}\boldsymbol{Z^T\widehat{W}ZG^{-1}\widehat{\beta}}_{ML}\\ &= \boldsymbol{G^{-1}\widehat{\beta}}_{ML}\end{aligned} \tag{18}$$

Estimator $\boldsymbol{\widehat{\gamma}}_{PR}$ didapatkan dengan menambahkan konstanta bias $k\mathbf{I}$, diperoleh:

$$\begin{aligned}\boldsymbol{\widehat{\gamma}}_{PR} &= (\Lambda_{PR} + k\mathbf{I})^{-1}\boldsymbol{Z^T\widehat{W}\hat{s}}\\ &= (\boldsymbol{B})^{-1}(\boldsymbol{XG})^{\boldsymbol{T}}\boldsymbol{\widehat{W}X\widehat{\beta}}_{ML}\\ &= (\boldsymbol{B})^{-1}\boldsymbol{G^T X^T \widehat{W} X G \widehat{\gamma}}_{ML}\\ &= (\boldsymbol{B})^{-1}\Lambda_{PR}\boldsymbol{\widehat{\gamma}}_{ML}\\ &= (\boldsymbol{B})^{-1}(\boldsymbol{B} - k\mathbf{I})\boldsymbol{\widehat{\gamma}}_{ML}\\ &= (\boldsymbol{B}^{-1}\boldsymbol{B} - k\mathbf{I}\boldsymbol{B}^{-1})\widehat{\gamma}_{ML}\\ &= (\mathbf{I} - k\boldsymbol{B}^{-1})\widehat{\gamma}_{ML}\end{aligned} \tag{19}$$

dengan $\boldsymbol{B} = \Lambda_{PRR} + k\mathbf{I} = \boldsymbol{Z^T\widehat{W}Z} \ + k\mathbf{I}$ dan $k > 0$

Untuk mendapatkan estimator JPR dimisalkan $\boldsymbol{\hat{s}}_{-i}, \boldsymbol{Z}_{-i}, \boldsymbol{\widehat{W}}_{-i}$ masing-masing menunjukkan vektor $\boldsymbol{s}$ dengan mengeluarkan baris ke-*i*, matriks $\boldsymbol{Z}$ dengan mengeluarkan baris ke-*i*, matriks $\boldsymbol{W}$ dengan mengeluarkan kedua baris dan kolom ke-*i*. Metode *jackknife* berperan dalam mereduksi bias sehingga dapat diperoleh penaksir parameter dengan bias yang kecil. Didefinisikan:

$$\boldsymbol{\widehat{\gamma}}_{PR-i} = \left(\boldsymbol{Z^T_{-i}\widehat{W}_{-i}Z_{-i}} + k\mathbf{B}\right)^{-1}\boldsymbol{Z^T_{-i}\widehat{W}_{-i}\hat{s}_{-i}} \tag{20}$$

Persamaan (19) merupakan estimator PR dengan mengeluarkan observasi ke-*i*, dimana: $\boldsymbol{Z^T_{-i}\widehat{W}_{-i}\hat{s}_{-i}} = \boldsymbol{Z^T\widehat{W}\hat{s}} - \boldsymbol{z_i^T\widehat{\mu}_i\hat{s}_i}$ dan $\left(\boldsymbol{Z^T_{-i}\widehat{W}_{-i}Z_{-i}} + k\mathbf{I}\right)^{-1}$ merupakan teorema dari *Sherman-Morrison Woodbury* menggunakan invers jumlah matriks yaitu:

$$\left(\boldsymbol{Z^T_{-i}\widehat{W}_{-i}Z_{-i}} + k\mathbf{I}\right)^{-1} = \left(\boldsymbol{Z^T\widehat{W}Z} + k\mathbf{I}\right)^{-1} + \frac{\left(\boldsymbol{Z^T\widehat{W}Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z_i z_i^T}\left(\boldsymbol{Z^T\widehat{W}Z} + k\mathbf{I}\right)^{-1}\boldsymbol{\widehat{\mu}_i}}{1 - \boldsymbol{z_i^T}\left(\boldsymbol{Z^T\widehat{W}Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z_i\widehat{\mu}_i}}$$

Sehingga Persamaan (20) menjadi:

$$\begin{aligned}\boldsymbol{\widehat{\gamma}}_{PR-i} &= \left(\boldsymbol{Z^T_{-i}\widehat{W}_{-i}Z_{-i}} + k\mathbf{I}\right)^{-1}\boldsymbol{Z^T_{-i}\widehat{W}_{-i}\hat{s}_{-i}}\\ &= \left[\left(\boldsymbol{Z^T\widehat{W}Z} + k\mathbf{I}\right)^{-1} + \frac{\left(\boldsymbol{Z^T\widehat{W}Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z_i z_i^T}\left(\boldsymbol{Z^T\widehat{W}Z}+k\mathbf{I}\right)^{-1}\boldsymbol{\widehat{\mu}_i}}{1-\boldsymbol{z_i^T}\left(\boldsymbol{Z^T\widehat{W}Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z_i\widehat{\mu}_i}}\right]\left(\boldsymbol{Z^T\widehat{W}\hat{s}} - \boldsymbol{z_i^T\widehat{\mu}_i\hat{s}_i}\right)\end{aligned}$$

$$\hat{\boldsymbol{\gamma}}_{PR-i} = \left[\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\hat{\boldsymbol{s}} + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\hat{\boldsymbol{s}}\right] - \left[\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i\right]$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\hat{\boldsymbol{s}}\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i} - \left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\hat{\boldsymbol{s}} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i - \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i} - \left[\frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}\right] - \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\hat{\boldsymbol{\mu}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i} - \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+kI\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}$$

$$= \hat{\boldsymbol{\gamma}}_{PR} - \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i(\hat{\boldsymbol{s}}_i-\boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR})}{1-\boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z}+k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i} \tag{21}$$

dengan $\hat{\boldsymbol{\gamma}}_{PR} = \left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\hat{\boldsymbol{s}}$

Menurut Hinkley (1977) mendefinisikan *weighted pseudo-value* sebagai

$$\boldsymbol{Q}_i = \hat{\boldsymbol{\gamma}}_{PR} + n(1-\boldsymbol{w}_i)(\hat{\boldsymbol{\gamma}}_{PR} - \hat{\boldsymbol{\gamma}}_{PR-i}) \tag{22}$$

Persamaan (21) disubtitusikan ke dalam Persamaan (22), sehingga diperoleh:

$$\boldsymbol{Q}_i = \hat{\boldsymbol{\gamma}}_{PR} + n(1-\boldsymbol{w}_i)\left(\hat{\boldsymbol{\gamma}}_{PR} - \hat{\boldsymbol{\gamma}}_{PR} + \frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\right)}{1 - \boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i}\right)$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + n(1-\boldsymbol{w}_i)\left(\frac{\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\right)}{1-\boldsymbol{w}_i}\right)$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + n\left(\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\right)\right) \tag{23}$$

dimana $\boldsymbol{w}_i = \boldsymbol{z}_i^T\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i$

Selanjutnya estimasi parameter dengan estimator JPR dilakukan dengan mengambil rata-rata *weighted pseudo-value.*

$$\hat{\boldsymbol{\gamma}}_{JPR} = \overline{\boldsymbol{Q}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Q}_i$$

$$= \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\gamma}}_{PR} + n\left(\left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\right)\right)$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + \left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\sum_{i=1}^{n}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR}\right)$$

$$= \hat{\boldsymbol{\gamma}}_{PR} + \left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\sum_{i=1}^{n}\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i - \left(\boldsymbol{Z}^T\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\sum_{i=1}^{n}(\boldsymbol{z}_i^T\hat{\boldsymbol{\mu}}_i\boldsymbol{z}_i^T\hat{\boldsymbol{\gamma}}_{PR})$$

$$\begin{aligned}\widehat{\boldsymbol{\gamma}}_{JPR} &= \widehat{\boldsymbol{\gamma}}_{PR} + \left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\widehat{\boldsymbol{s}} - \left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z}\widehat{\boldsymbol{\gamma}}_{PR} \\ &= \widehat{\boldsymbol{\gamma}}_{PR} + \widehat{\boldsymbol{\gamma}}_{PR} - \left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z}\widehat{\boldsymbol{\gamma}}_{PR} \\ &= \widehat{\boldsymbol{\gamma}}_{PR} + \left(\mathbf{I} - \left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z}\right)\widehat{\boldsymbol{\gamma}}_{PR} \\ &= \widehat{\boldsymbol{\gamma}}_{PR} + (k\left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1})\widehat{\boldsymbol{\gamma}}_{PR} \\ &= \left(\mathbf{I} + k\left(\boldsymbol{Z}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{Z} + k\mathbf{I}\right)^{-1}\right)\widehat{\boldsymbol{\gamma}}_{PR} \\ &= (\mathbf{I} + k(\boldsymbol{B})^{-1})\widehat{\boldsymbol{\gamma}}_{PR} \end{aligned} \tag{24}$$

Estimator *Modified Jackknifed Poisson Ridge* dapat dituliskan sebagai berikut:

$$\widehat{\boldsymbol{\gamma}}_{MJPR} = (\mathbf{I} - k^2\boldsymbol{B}^{-2})(\boldsymbol{I} - k\boldsymbol{B}^{-1})\hat{\gamma}_{ML} \tag{25}$$

1.5.8 Pemilihan Nilai k

Masalah utama yang dihadapi dalam mengatasi multikolinearitas dalam menggunakan analisis regresi *ridge* adalah penentuan nilai k. Proses pemilihan nilai k ini sangat bergantung pada studi kasus yang sedang ditangani. Namun, tidak ada aturan pasti untuk memilih nilai *k* dalam regresi *ridge*, sehingga terdapat beberapa nilai *ridge* k yang telah dikemukakan oleh peneliti sebelumnya. Salah satu metode yang dapat digunakan dan umum digunakan untuk estimasi MJPR adalah yang diperkenalkan oleh Schaeffer et.al (1984), yaitu:

$$k = \frac{1}{\hat{\alpha}^2{}_{max}} \tag{26}$$

dengan $\hat{\alpha}^2{}_{max}$ adalah nilai maksimum dari $\left(\boldsymbol{G}^{\boldsymbol{T}}\widehat{\boldsymbol{\beta}}^*_{ML}\right)^2$

1.5.9 Uji Kebaikan Model

Pemilihan model terbaik pada analisis regresi dapat dilihat menggunakan beberapa metode salah satunya melalui nilai *Akaike Information Criteria* (AIC). Kriteria pemilihan model terbaik berdasarkan nilai AIC dilihat dari nilai AIC yang terkecil. Sehingga jika terdapat nilai $AIC_A < AIC_B$, maka model A dianggap sebagai model terbaik. Nilai AIC dapat didefinisikan sebagai (Safira, 2023):

$$AIC = -2ln\left(l(\hat{\beta})\right) + 2c \tag{27}$$

dengan:

$l(\hat{\beta})$: nilai log likelihood dari model

c : jumlah parameter

1.5.10 Uji Signifikansi Parameter

Pengujian signifikansi parameter dalam model regresi dilakukan untuk menentukan hubungan antara variabel respon dan variabel prediktor.

1. Uji Simultan (Uji F)

Pengujian serentak atau pengujian simultan pada model GPR dengan estimasi MJPR dilakukan untuk mengetahui apakah variabel prediktor secara bersama-sama berpengaruh terhadap variabel respon. Hipotesis yang digunakan dalam pengujian ini adalah:

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$ (seluruh variabel prediktor secara simultan tidak berpengaruh terhadap variabel respon)

$H_0 : min\ ada\ \beta_j \neq 0;\ j = 1, 2, \ldots, p$ (seluruh variabel prediktor secara simultan berpengaruh terhadap variabel respon)

Statistik uji yang digunakan yaitu:

$$F = \frac{JKR/p}{JKG/(n-p-1)} = \frac{KTR}{KTG} \tag{28}$$

dengan kriteria pengujian hipotesis yaitu tolak H_0 jika nilai $F > F_{(\alpha,p,n-p-1)}$ atau jika $p - value < \alpha = 0.05$ yang berarti bahwa ada salah satu parameter yang berpengaruh secara signifikan.

2. Uji Parsial (Uji T)

Pengujian parsial pada model GPR dengan estimasi MJPR dilakukan untuk menilai pengaruh masing-masing variabel prediktor terhadap variabel respon secara individu.

Hipotesis yang digunakan adalah:

$H_0 : \beta_j = 0; j = 1, 2, \ldots, p$ (masing-masing variabel prediktor tidak berpengaruh signifikan terhadap variabel respon)

$H_0 : \beta_j \neq 0; j = 1, 2, \ldots, p$ (masing-masing variabel prediktor berpengaruh signifikan terhadap variabel respon)

Statistik uji yang digunakan yaitu:

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{29}$$

dengan $\hat{\beta}_j$ merupakan estimasi parameter variabel ke-j dan $SE(\hat{\beta}_j)$ adalah standard *error* variabel ke-j. Adapun kriteria pengujian hipotesis yaitu Tolak H_0 jika nilai $|t_{hit}| > t_{(\alpha/2pn-p-1)}$ atau jika $p - value < \alpha = 0.05$ (Arrasyid, dkk., 2021).

1.5.11 Angka kematian ibu

Angka kematian ibu adalah banyaknya perempuan yang meninggal selama masa kehamilan, persalinan dan masa nifas (42 hari setelah melahirkan) tanpa memperhitungkan lama kehamilan per 100.000 kelahiran hidup, yang meninggal karena penyebab kematian terkait dengan gangguan kehamilan atau penanganannya, namun tidak termasuk kecelakaan atau kasus insedentil (Dinkes Sulsel, 2021).

Untuk mengurangi angka kematian ibu, Kementerian Kesehatan Republik Indonesia (2021) melakukan beberapa upaya yang terdiri:

1. Pemberian imunisasi tetanus yang diberikan kepada wanita usia subur, termasuk ibu hamil, sebagai langkah preventif terhadap infeksi yang seringkali menjadi penyebab kematian ibu. Imunisasi Td ditargetkan pada wanita usia subur yang hamil atau tidak, dalam kelompok usia 15-39 tahun. Untuk ibu hamil, dikatakan bahwa mereka telah menerima imunisasi Td2+ jika telah mendapatkan imunisasi Td2 hingga Td5.

2. Pemberian tablet tambah darah pada ibu hamil dilakukan untuk mencegah anemia yang dapat meningkatkan risiko kelahiran prematur, kematian ibu dan anak, serta penyakit infeksi.
3. Pelayanan kesehatan ibu nifas dengan pemberian kapsul vitamin A pada ibu nifas yang dilakukan minimal empat kali dengan waktu kunjungan ibu dan bayi baru lahir bersamaan, yaitu pada enam jam sampai dengan dua hari setelah persalinan, pada hari ketiga sampai hari ke tujuh, pada hari ke delapan sampai hari ke 28, dan pada hari ke 29 sampai dengan 42 hari setelah persalinan.
4. Pelayanan kontrasepsi, yang mencakup konseling, pemberian, pemasangan, serta pencabutan kontrasepsi dalam rahim, diberikan untuk mencegah kehamilan dan menangani komplikasi atau efek samping yang mungkin terjadi. KB merupakan salah satu strategi untuk menciptakan lingkungan yang mendukung kesehatan ibu dan mengurangi angka kematian ibu secara signifikan.
5. Pelayanan kesehatan ibu bersalin oleh tenaga kesehatan yang kompeten dan terlatih yang dilakukan dengan bantuan medis seperti dokter, perawat maupun bidan menjadi salah satu upaya dalam menurunkan angka kematian ibu yang dilakukan di fasilitas pelayanan kesehatan. Ketersediaan tenaga kesehatan dalam intitusi rumah sakit atau puskesmas pada suatu daerah sangatlah penting karena jika jumlah tenaga medis terpenuhi dalam suatu daerah maka pelayanan kesehatan akan terjamin.

BAB II
METODOLOGI PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan data kuantitatif mengenai Jumlah Kematian Ibu di Provinsi Sulawesi Selatan pada Tahun 2021. Data tersebut diperoleh dari buku Profil Kesehatan Sulawesi Selatan tahun 2022 yang diterbitkan oleh Dinas Kesehatan Provisi Sulawesi Selatan. Unit penelitian yang digunakan dalam penelitian ini adalah 24 kabupaten/kota Provinsi Sulawesi Selatan yang terdiri dari 21 Kabupaten dan 3 Kota.

2.2 Variabel Penelitian

Dalam penelitian ini, variabel yang digunakan terdiri dari 1 variabel dependen (Y) dan 5 variabel independen (X) dengan rincian sebagai berikut:

Tabel 2. Variabel Penelitian

No.	Variabel	Keterangan
1.	Y	Jumlah Kematian Ibu
2.	X_1	Jumlah Ibu Hamil yang Mendapatkan Tablet Tambah Darah
3.	X_2	Jumlah Cakupan Imunisasi Td2+ pada Ibu Hamil
4.	X_3	Jumlah Pasangan Usia Subur yang Melakukan KB Aktif
5.	X_4	Perkiraan Jumlah Ibu Hamil dengan Komplikasi Kebidanan
6.	X_5	Jumlah Tenaga keperawatan dan Kebidanan di Fasilitas Kesehatan

2.3 Tahapan Analisis

Tahapan analisis yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Melakukan input data yang terdiri dari 1 variabel respon dan 5 variabel prediktor (Lampiran 1).
2. Melakukan analisis data pada data penelitian
   a. Menyajikan statistik deskriptif dari variabel respon dan prediktor.
   b. Melakukan uji distribusi Poisson menggunakan statistik uji *Kolmogorov-Smirnov.*
3. Melakukan pengujian equidisversi dengan menghitung nilai *deviance* dan *pearson chi-square yang* dibagi dengan derajat bebas.
4. Melakukan pendeteksian multikolinearitas dengan melihat nilai VIF dan korelasi antar variabel prediktornya.

5. Memodelkan data menggunakan *Generalized Poisson Regression* dengan estimasi *Modifed Jackknifed Poisson Ridge* untuk mengatasi multikolinieritas dengan langkah-langkah sebagai berikut:
   a. Melakukan transformasi korelasi dengan metode *centering and scaling* terhadap variabel prediktor menggunakan Persamaan (16).
   b. Menganalisis model regresi *generalized Poisson regression* dengan *maximum likelihood estimation* dari hasil data *centering and scaling* ($\widehat{\boldsymbol{\beta}}^*_{ML}$).
   c. Membentuk Matriks $\boldsymbol{X}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{X}$ dengan data hasil transformasi, dimana $\widehat{\boldsymbol{W}}$ adalah matriks diagonal dari $\boldsymbol{\mu} = (\boldsymbol{X}\widehat{\boldsymbol{\beta}}^*_{ML})$.
   d. Membentuk matriks $\boldsymbol{Z} = \boldsymbol{XG}$ dimana $\boldsymbol{G}$ yang merupakan *vector eigen* dari matriks $\boldsymbol{X}^{\boldsymbol{T}}\widehat{\boldsymbol{W}}\boldsymbol{X}$.
   e. Menghitung estimator $\widehat{\boldsymbol{\gamma}}^*_{\boldsymbol{ML}} = \boldsymbol{G}^{-1}\widehat{\boldsymbol{\beta}}^*_{ML}$.
   f. Mencari nilai tetapan bias *k* menggunakan Persamaan (26).
   g. Mensubtitusikan nilai *k* untuk mendapat estimator *modified jackknifed Poisson ridge* menggunakan Persamaan (25).
   h. Membentuk model *generalized Poisson regression* dengan estimasi *modified jackknifed Poisson ridge* dari data hasil transformasi dengan mensubstitusi $\widehat{\boldsymbol{\gamma}}_{MJPR}$ sehingga didapatkan persamaan regresi *ridge* yang asli dengan transformasi $\widehat{\boldsymbol{\beta}}^*_{MJPR} = \boldsymbol{G}\widehat{\boldsymbol{\gamma}}^*_{MJPR}$.
   i. Mentransformasikan persamaan regresi ke bentuk awal tanpa *centering* dan *scaling* dan membentuk persamaan model *generalized Poisson regression* dengan estimasi *modified jackknifed Poisson ridge* yang baru dengan rumus yaitu:
   $$\widehat{\boldsymbol{\beta}}_i = \left(\frac{1}{\sqrt{n-1}S_{ij}}\right)\widehat{\boldsymbol{\beta}}^*{}_{MJPR}$$
   $$\hat{\beta}_0 = \text{-} -\hat{\beta}_{MJPR}\bar{X}_1 - -\hat{\beta}_{MJPR}\bar{X}_2 - \ldots -\hat{\beta}_{MJPRp}\bar{X}_p$$
   dengan S_j merupakan standar deviasi data awal X_j, sedangkan $\bar{X}_j$ adalah rata-rata data awal X_j sebelum ditransformasi.
   j. Pengujian Signifikansi Parameter model *Generalized Poisson Regression* dengan estimasi *Modified Jackknifed Poisson Ridge.*
   k. Uji kebaikan model
   l. Interpretasi dan menarik kesimpulan.