PERBANDINGAN PERFORMA *JAMES-STEIN ESTIMATOR, RIDGE REGRESSION ESTIMATOR*, DAN *MODIFIED KIBRIA-LUKMAN ESTIMATOR* DALAM MENGATASI MULTIKOLINEARITAS PADA REGRESI POISSON: SIMULASI STUDI

(Skripsi)

Oleh

M. FIKRI ALYASA ZAM ZAMI

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVESITAS LAMPUNG
BANDAR LAMPUNG
2025

ABSTRACT

COMPARISON THE PERFORMANCE OF THE JAMES-STEIN ESTIMATOR, RIDGE REGRESSION ESTIMATOR AND MODIFIED KIBRIA LUKMAN ESTIMATOR IN OVERCOMING MULTICOLLINEARITY IN POISSON REGRESSION: SIMULATION STUDY

BY

M. FIKRI ALYASA ZAM ZAMI

Poisson regression is a statistical method used to analyze data with a response in the form of a count variable. This regression uses the Maximum Likelihood Estimation (MLE) method to estimate model parameters. The purpose of this study is to compare the performance of the Poisson James-Stein Estimator (PJSE), Poisson Ridge Regression Estimator (PRRE), and Poisson Modified Kibria-Lukman Estimator (PMKLE) methods in dealing with multicollinearity using simulated data with n = 20, 40, 60 and 80 in poisson model (p=6) with ρ = 0.3 and 0.99. The best model was compered based on the MSE value. The results showed that in the partial correlation and full correlation data of the PRRE method of k_2 parameters and PMKLE parameters k_2 were better in overcoming multicollinearity.

Keywords: James-Stein Estimator, Ridge Regression Estimator, Modified Kibria-Lukman Estimator, Multicollinearity.

ABSTRAK

PERBANDINGAN PERFORMA *JAMES-STEIN ESTIMATOR, RIDGE REGRESSION ESTIMATOR,* DAN *MODIFIED KIBRIA-LUKMAN ESTIMATOR* DALAM MENGATASI MULTIKOLINEARITAS PADA REGRESI POISSON: SIMULASI STUDI

Oleh

M. FIKRI ALYASA ZAM ZAMI

Regresi Poisson merupakan statistika yang digunakan untuk menganalisis data dengan respon berupa variabel hitungan. Regresi ini menggunakan metode *Maximum Likelihood Estimation* (MLE) untuk mengestimasi parameter model. Tujuan dari penelitian ini adalah untuk membandingkan kinerja metode *Poisson James-Stein Estimator* (PJSE)*, Poisson Ridge Regression Estimator* (PRRE), dan *Poisson Modified Kibria-Lukman Estimator* (PMKLE) dalam mengatasi multikolinearitas menggunakan data simulasi dengan n = 20, 40, 60 dan 80 pada model Poisson (p=6) dengan ρ = 0.3 dan 0.99. Model terbaik dibandingkan berdasarkan nilai MSE. Hasil penelitian menunjukkan bahwa pada data korelasi partial dan korelasi penuh metode PRRE parameter k_2 dan PMKLE parameter k_2 lebih baik dalam mengatasi multikolinearitas.

Keywords: *James-Stein Estimator*, *Ridge Regression Estimator*, *Modified Kibria-Lukman Estimator*, Multikolinaritas.

PERBANDINGAN PERFORMA *JAMES-STEIN ESTIMETOR, RIDGE REGRESSION ESTIMATOR,* DAN *MODIFIED KIBRIA-LUKMAN ESTIMATOR* DALAM MENGATASI MULTIKOLINEARITAS PADA REGRESI POISSON: SIMULASI STUDI

Oleh

M. Fikri Alyasa Zam Zami

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVESITAS LAMPUNG
BANDAR LAMPUNG
2025

Judul : **PERBANDINGAN PERFORMA *JAMES-STEIN ESTIMATOR, RIDGE REGRESSION ESTIMATOR,* DAN *MODIFIED KIBRIA-LUKMAN ESTIMATOR* UNTUK MENGATASI MULTIKOLINEARITAS PADA REGRESI POISSON: SIMULASI STUDI**

Nama Mahasiswa : *M. Fikri Alyasa Zam Zami*

NPM : **2117031016**

Program Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI,

1. Komisi Pembimbing

Prof. Ir. Netti Herawati, M.Sc., Ph.D.	**Misgiyati, S. Pd., M. Si.**
NIP. 196501251990032001	NIP. 198509282023212032

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Prof. Ir. Netti Herawati, M.Sc., Ph.D.**	....
Sekretaris	: **Misgiyati, S. Pd., M. Si.**	....
Penguji Bukan Pembimbing	: **Drs. Nusyirwan, M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP 197110012005011002

Tanggal Lulus Ujian Skripsi: **24 Maret 2025**

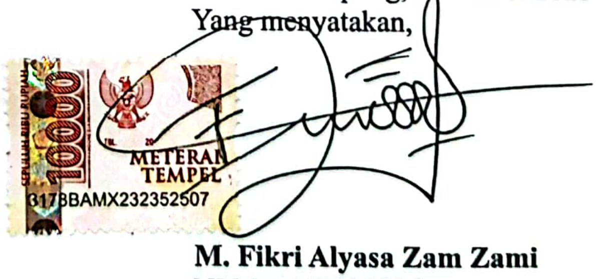
PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama	: **M. Fikri Alyasa Zam Zami**
Nomor Pokok Mahasiswa	: **2117031016**
Jurusan	: **Matematika**
Judul Skripsi	: **PERBANDINGAN PERFORMA *JAMES-STEIN ESTIMETOR, RIDGE REGRESSION ESTIMATOR*, DAN *MODIFIED KIBRIA-LUKMAN ESTIMATOR* DALAM MENGATASI MULTIKOLINEARITAS PADA REGRESI POISSON: SIMULASI STUDI**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku

Bandar Lampung, 24 Maret 2025
Yang menyatakan,



M. Fikri Alyasa Zam Zami
NPM. 2117031016

RIWAYAT HIDUP

Penulis yang bernama lengkap M. Fikri Alyasa Zam Zami, lahir di Kalianda pada 8 Agustus 2002. Penulis lahir dari pasangan Bapak M. Amir dan Ibu Rostina Idawati dan merupakan anak ketiga dari empat bersaudara. Penulis saat ini bertempat tinggal di Desa Maja, Kecamatan Kalianda, Kabupaten Lampung Selatan, Lampung. Penulis menempuh pendidikan sekolah dasar di MIN 1 Lampung Selatan tahun 2009-2015. Penulis lalu menempuh jenjang sekolah menengah pertama di MTs Negeri 1 Lampung Selatan tahun 2015-2018 dan selanjutnya ke jenjang sekolah atas di SMAN 1 Kalianda tahun 2018-2021.

Pada tahun 2021 penulis terdaftar sebagai Mahasiswa Program Studi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswa, penulis aktif mengikuti organisasi kemahasiswaan yakni menjadi pengurus HIMATIKA FMIPA Universitas Lampung sebagai anggota Biro Dana dan Usaha periode 2022.

Pada bulan Desember 2023 - Februari 2024, penulis melaksanakan Kerja Praktik di BPS Kabupaten Lampung Selatan sebagai bentuk penerapan ilmu yang telah diperoleh selama kuliah. Pada bulan Juni - Agustus 2024, penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Negeri Jemanten, Kecamatan Marga Tiga, Kabupaten Lampung Timur sebagai bentuk pengabdian Mahasiswa dan menjalankan Tri Darma Perguruan Tinggi.

KATA INSPIRASI

"*The winner takes it all and the loser has to fall"*

(ABBA)

PERSEMBAHAN

Alhamdulillah, puji syukur kepada Allah SWT. Yang telah memberikan petunjuk dan rahmat-Nya juga memberikan penerangan dalam ilmu pengetahuan. Hanya karenanya-Nya lah skripsi ini bisa penulis selesaikan dengan rasa syukur dan bahagia. Dengan segala kerendahan hati, penulis persembahkan karya sederhana ini kepada:

Orang Tua dan Saudara

Yang telah memberikan dukungan, doa, serta kasih sayangnya sehingga menjadi sumber kekuatan penulis untuk bertahan dan berjuang menyelesaikan skripsi ini. Sebuah perjuangan yang panjang dan penuh cerita yang tidak lepas dari doa tulus Ayah, Ibu, Kakak, Abang, Adik dan seluruh keluarga besar.

Dosen Pembimbing dan Pembahas

Yang senantiasa memberikan bimbingan, arahan, dan ilmu yang sangat bermanfaat bagi penulis.

Almamaterku Tercinta, Universitas Lampung

SANWACANA

Segala puji bagi Allah, Tuhan semesta alam, atas limpahan karunia dan rahmat Nya lah, penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Performa *James-Stein Estimetor*, *Ridge Regression Estimator*, dan *Modified Kibria-Lukman Estimator* Dalam Mengatasi Multikolinearitas Pada Regresi Poisson: Simulasi Studi”.

Penulis menyadari bahwa dalam penyelesaian skripsi ini, tidak lepas dari dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Ibu Prof. Ir. Netti Herawati, M.Sc., Ph.D., selaku dosen pembimbing satu sekaligus sebagai dosen pembimbing akademik atas bimbingan, nasihat, arahan, dan motivasi yang membangun kepada penulis selama proses penyusunan skripsi ini.
2. Ibu Misgiyati, S. Pd., M. Si., selaku dosen pembimbing dua yang telah memberikan bimbingan serta kemudahan dalam penyusunan skripsi ini.
3. Bapak Drs. Nusyirwan, M.Si., selaku dosen penguji yang telah memberikan evaluasi dan saran bagi perbaikan skripsi penulis.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh dosen dan staff Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Orang tua dan seluruh saudara tercinta, Ayah, Ibu, Kakak, Abang, Adik, dan seluruh keluarga besar lainnya yang telah dengan sabar memberikan segala

dukungan dan kepercayaan selama ini.
8. Semua pihak yang tidak bisa penulis sebutkan satu persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi penulis berharap semoga skripsi ini dapat bermanfaat bagi kita semua.

Bandar Lampung, 24 Maret 2025
Penulis,

M. Fikri Alyasa Zam Zami

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Regresi Poisson adalah salah satu metode statistik yang digunakan untuk menganalisis data dengan respon berupa variabel *count* atau hitungan. Namun, dalam penerapannya, seringkali terdapat masalah multikolinearitas, yaitu situasi di mana dua atau lebih variabel independen memiliki korelasi tinggi. Multikolinearitas dapat menjadi masalah dalam mempengaruhi kesimpulan mengenai signifikansi dan estimasi parameter. Tingkat multikolinearitas yang tinggi dapat menyebabkan varians yang besar dalam estimasi kuadrat terkecil dari koefisien beta dalam model regresi, serta dapat menghasilkan hasil yang bias (Lavery *et al*., 2019).

Untuk mengatasi masalah multikolinearitas pada regresi Poisson, beberapa estimator alternatif telah diusulkan oleh para peneliti. Diantaranya adalah *Poisson Ridge Regression Estimator* (PRE) yang diperkenalkan oleh Månsson dan Shukur (2011), *Poisson Modified Kibria-Lukman Estimator* (PMKLE) oleh Aladeitan *et al.* (2021) dan juga *Poisson James-Stein estimator* (PJSE) oleh Amin *et al.* (2020).

James-Stein Estimator adalah metode yang bisa membantu mengurangi varians dalam perkiraan koefisien regresi, terutama saat ada masalah multikolinearitas atau banyaknya variabel independen dalam model. Metode ini menggunakan teknik penyusutan (*shrinkage*) untuk mengurangi ketidakstabilan dalam estimasi akibat multikolinearitas. Dengan pendekatan ini, estimasi koefisien yang dihasilkan biasanya lebih baik dibandingkan metode *Maximum Likelihood*

Estimation (MLE) (Judge *et al*., 1985).

Poisson Ridge Regression adalah bentuk pengembangan dari regresi Poisson yang memadukan pendekatan regresi Poisson dengan teknik *ridge regression* untuk mengatasi masalah multikolinearitas. Metode ini dikembangkan oleh Mansson dan Shukur pada tahun 2011 sebagai hasil modifikasi dari regresi Ridge, yang pertama kali diperkenalkan oleh Hoerl dan Kennard pada tahun 1970 untuk menangani multikolinearitas.

M. Golam Kibria dan M.A. Lukman memperkenalkan *Kibria-Lukman Estimator* sebagai solusi untuk masalah multikolinearitas dalam model regresi Poisson. Tujuan utama metode ini adalah meningkatkan akurasi estimasi parameter dibandingkan dengan metode *Ridge Regression* yang sudah ada. Selanjutnya, pada tahun 2021, Aladeitan dan rekan-rekannya mengembangkan versi modifikasi dari *estimator* ini, yang dikenal sebagai *Modified Kibria-Lukman Estimator*. Modifikasi ini dirancang khusus untuk menangani multikolinearitas dalam model regresi Poisson, dengan fokus pada peningkatan efisiensi estimasi parameter. Hal ini menjadi sangat penting ketika variabel prediktor memiliki korelasi tinggi, yang dapat mempengaruhi stabilitas dan akurasi hasil estimasi.

Dalam penelitian sebelumnya, metode *Poisson James-Stein Estimator* telah dikaji untuk menyelesaikan masalah multikolinearitas pada model regresi Poisson oleh Amin *et al.* (2020) yang menggunakan simulasi Monte Carlo dan juga data tentang kerusakan pesawat untuk mengevaluasi kinerja *James-Stein Estimator*. Hasilnya menunjukkan bahwa *Poisson James-Stein Estimator* menghasilkan nilai *Mean Square Error* (MSE) yang lebih rendah dibandingkan dengan *Maximum Likelihood Estimation* (MLE) dan metode lain. Selain itu, penelitian Oghenekevwe *et al*. (2021) mengkaji *Poisson Ridge Regression Estimator* menggunakan data simulasi. Hasilnya menunjukkan bahwa metode ini efektif untuk menangani multikolinearitas dalam model regresi Poisson berdasarkan parameter ridge *k* yang digunakan. Lalu *Poisson Modified Kibria-Lukman Estimator* (PMKLE) dikaji oleh Aladeitan *et al.* pada tahun 2021 menggunakan data simulasi dan studi kasus untuk mengevaluasi kinerja PMKLE dalam menghadapi multikolinearitas dan terbukti lebih efisien dibandingkan *estimator* lainnya. Meskipun studi-studi sebelumnya telah membandingkan beberapa *estimator* tersebut secara terpisah, belum ada penelitian yang secara langsung membandingkan performa ketiga *estimator* tersebut.

Oleh karena itu, penelitian ini bertujuan untuk melakukan perbandingan antara PJSE, PRRE, dan PMKLE dalam mengatasi multikolinearitas pada model regresi Poisson. Selain itu, penelitian ini juga akan menggunakan skenario simulasi. Hal ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang kinerja dari masing-masing *estimator* dalam berbagai kondisi.

Dengan melakukan perbandingan ini, penelitian diharapkan dapat memberikan kontribusi signifikan dalam pemilihan *estimator* yang optimal untuk mengatasi multikolinearitas pada regresi Poisson.

1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis dan membandingkan performa *Poisson James-Stein Estimator* (PJSE), *Poisson Ridge Regression Estimator* (PRRE) dan *Poisson Modified Kibria-Lukman Estimator* (PMKLE) dalam mengatasi multikolinearitas pada regresi Poisson.
2. Menentukan metode mana yang lebih efektif dalam menghasilkan estimasi parameter yang stabil dan akurat ketika multikolinearitas tinggi terjadi.

1.3 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penelitian ini adalah untuk memperluas wawasan bagi penulis dan memberikan wawasan kepada pembaca dan peneliti terkait perbandingan performa *Poisson James-Stein Estimator* (PJSE), *Poisson Ridge Regression Estimator* (PRRE) dan *Poisson Modified Kibria-Lukman Estimator* (PMKLE) dalam mengatasi multikolinearitas pada regresi Poisson.

II. TINJAUN PUSTAKA

2.1 Distribusi Poisson

Distribusi Poisson adalah distribusi probabilitas berapa kali suatu peristiwa acak terjadi. Distribusi ini berkaitan dengan nilai-nilai pada variabel acak X, di mana jumlah kejadian yang terjadi dalam suatu periode tertentu dinyatakan sebagai peluang jika rata-rata kejadian tersebut diketahui dan terjadi secara independen dalam waktu yang berbeda. Distribusi Poisson mempunyai karakteristik yaitu peubah acaknya diskrit dan informasi mengenai besarnya nilai rata-rata dari suatu kejadian dalam suatu interval waktu tertentu (Pudjoatmodjo & Hendayun, 2016).

Menurut Walpole (1995), fungsi peluang peubah acak Y dengan parameter μ pada sebaran Poisson dapat dirumuskan sebagai berikut:

$$p(y, \mu) = \frac{e^{-\mu}\mu^{y}}{y!}; = 0, 1, 2, 3, \ldots; i = 1, 2, 3, \ldots, n \qquad (2.1)$$

dengan:

$p(y, \mu)$: fungsi peluang dari distribusi Poisson,

y : nilai aktual jumlah kejadian (0, 1, 2, …),

μ : rata – rata jumlah kejadian dalam selang waktu tertentu,

e : bilangan Euler (2.7183).

Pada distribusi Poisson memiliki nilai *mean* dan varian yang sama yaitu $E(Y)$ = Var (Y) = μ. Hal ini dapat dibuktikan melalui penjabaran berikut:

$$
\begin{aligned}
E(Y) &= \sum_{y=0}^{\infty} y \cdot p(y,\mu) \\
&= \sum_{y=0}^{\infty} y \frac{e^{-\mu}\mu^{y}}{y!} \\
&= \sum_{y=0}^{1} y \cdot \frac{e^{-\mu}\mu^{y}}{y(y-1)!} + \sum_{y=1}^{\infty} y \cdot \frac{e^{-\mu}\mu^{y}}{y(y-1)!} \\
&= 0 + \sum_{y=1}^{\infty} y \frac{e^{-\mu}\mu^{y}}{y(y-1)!} \\
&= \sum_{y=1}^{\infty} \frac{e^{-\mu}\mu \cdot \mu^{(y-1)}}{(y-1)!} \\
&= \mu \sum_{z=0}^{\infty} \frac{e^{-\mu}\mu^{z}}{(z)!} \qquad \text{(Misalkan z = y − 1 dan y = 1)} \\
&= \mu \cdot 1 \\
&= \mu.
\end{aligned}
$$

$$
\begin{aligned}
Var(Y) &= E(Y^2) - [E(Y)]^2 \\
&= E(Y^2 - Y + Y) - [E(Y)]^2 \\
&= E[Y(Y-1) + Y] - [E(Y)]^2 \\
&= E[Y(Y-1)] + E[Y] - [E(Y)]^2 \\
&= \left(\sum_{y=0}^{\infty} y(y-1) \frac{e^{-\mu}\mu^{y}}{y!} \right) + \mu - \mu^2 \\
&= \left(\sum_{y=0}^{\infty} y(y-1) \frac{e^{-\mu}\mu^{y}}{y(y-1)(y-2)!} \right) + \mu - \mu^2 \\
&= \left(\sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^{y}}{(y-2)!} \right) + \mu - \mu^2
\end{aligned}
$$

$$= \left(\mu^2 \sum_{y=2}^{\infty} \frac{e^{-\mu}\, \mu^{y-2}}{(y-2)!} \right) + \mu - \mu^2$$

$$= \mu^2 \sum_{z=0}^{\infty} \frac{e^{-\mu}\, \mu^{z}}{z!} + \mu - \mu^2 \qquad (\text{Misalkan z} = \text{y} - 2 \text{ dan y} = 2)$$

$$= \mu^2 + \mu - \mu^2$$

$$= \mu.$$

2.2 Regresi Poisson

Regresi Poisson adalah teknik regresi nonlinier yang digunakan untuk memodelkan hubungan antara variabel dependen yang berbentuk variabel acak diskrit dengan distribusi Poisson dan variabel independen. Model regresi Poisson adalah sebagai berikut:

$$y_i = \mu_i + \varepsilon_i \qquad (2.2).$$

Selanjutnya, dalam regresi Poisson hubungan tersebut dapat dituliskan dalam bentuk:

$$E(y_i) = \mu_i = exp\big(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}\big) = exp(x_i^T \beta) \qquad (2.3)$$

dengan:

y : jumlah kejadian,

ε_i : sisaan ke-i,

B : koefisien regresi Poisson.

Dalam model regresi Poisson, kita menggunakan metode *Maximum Likelihood Estimation* (MLE) untuk menaksir parameter dengan cara memaksimumkan fungsi *likelihood* (Myers, 1990). Fungsi *likelihood* untuk model regresi Poisson dapat dinyatakan sebagai berikut:

$$L(y;\mu) = \prod_{i=1}^{n} P(y_i;\mu)$$

$$= \prod_{i=1}^{n} \frac{e^{-\mu_i}\,\mu_i^{yi}}{y_i!}$$

$$= \frac{\prod_{i=1}^{n} e^{-\mu_i}\,\mu_i^{yi}}{\prod_{i=1}^{n} y_i!}$$

$$= \frac{\{(\prod_{i=1}^{n} e^{-\mu_i})(\prod_{i=1}^{n} \mu_i^{yi})\}}{\prod_{i=1}^{n} y_i!}$$

$$= \frac{\{(e^{-\sum_{i=1}^{n}\mu_i})(\prod_{i=1}^{n} \mu_i^{yi})\}}{\prod_{i=1}^{n} y_i!} \qquad (2.4).$$

Untuk memudahkan penyelesaian persamaan di atas, maka dibentuk fungsi *log likelihood* seperti berikut:

$$logL(y;\mu) = log\,\frac{\{(e^{-\sum_{i=1}^{n}\mu_i})(\prod_{i=1}^{n} \mu_i^{yi})\}}{\prod_{i=1}^{n} y_i!}$$

$$= \left\{log\left(e^{-\sum_{i=1}^{n}\mu_i} + \prod_{i=1}^{n} \mu_i^{yi}\right)\right\} - \left\{log\prod_{i=1}^{n} y_i!\right\}$$

$$= \left\{\left(log\; e^{-\sum_{i=1}^{n}\mu_i}\right) + \left(log\prod_{i=1}^{n} \mu_i^{yi}\right)\right\} - \left\{log\prod_{i=1}^{n} y_i!\right\}$$

$$= -\sum_{i=1}^{n}\mu_i + \sum_{i=1}^{n} y_i\; log\; \mu_i - \sum_{i=1}^{n}\; log\; y_i!$$

$$= \sum_{i=1}^{n} y_i\; log\; \mu_i - \sum_{i=1}^{n}\mu_i - \sum_{i=1}^{n}\; log\; y_i!$$

$$= \sum_{i=1}^{n} y_i(x_i^T\beta) - \sum_{i=1}^{n}\; exp(x_i^T\beta) - \sum_{i=1}^{n}\; log\; y_i! \qquad (2.5).$$

Untuk mendapatkan nilai *estimator maximum likelihood* $\hat{\beta}$, vektor koefisien diestimasi menggunakan metode *maximum likelihood* dengan cara mengambil turunan pertama dari log $L(y;\beta)$ terhadap β sebagai berikut:

$$S(\beta) = \frac{\partial \; log \; L(y;\beta)}{\partial \beta} = 0$$

$$\frac{\partial \left\{\sum_{i=1}^{n} y_i(x_i^T \beta) - \sum_{i=1}^{n} e^{(x_i^T \beta)} - \sum_{i=1}^{n} \; log \; y_i!\right\}}{\partial \beta} = 0$$

$$\sum_{i=1}^{n} y_i(x_i^T) - (x_i^T) \sum_{i=1}^{n} e^{(x_i^T \beta)} = 0$$

$$\sum_{i=1}^{n} \left(y_i - e^{(x_i^T \beta)}\right)(x_i^T) = 0 \qquad (2.6).$$

Pada persamaan di atas, terdapat persamaan non-linier untuk parameter β. Selanjutnya digunakan algoritma *Iterative Weighted Least Square* (IWLS) untuk menemukan solusi dari nilai vektor $S(\beta)$. IWLS merupakan pengembangan dari metode *Fisher Scoring*. Berdasarkan fungsi penghubung log(μi) = ηi = $x_i^T \beta$, iterasi *Fisher Scoring* dilakukan sebagai berikut:

$$\begin{aligned}
\beta^t &= \beta^{(t-1)} + [X^T W X]^{-1} X^T A (Y - \mu) \\
&= [X^T W X]^{-1} X^T W X \beta^{(t-1)} + [X^T W X]^{-1} X^T A (Y - \mu) \\
&= \left[X^T \widehat{W} X\right]^{-1} \left[X^T W X \beta^{(t-1)} + X^T A (Y - \mu)\right] \\
&= [X^T W X]^{-1} \left[X^T W X \beta^{(t-1)} + X^T W \left(\frac{\partial \eta}{\partial \mu}\right)(Y - \mu)\right] \\
&= [X^T W X]^{-1} X^T W \left[X \beta^{(t-1)} + \left(\frac{\partial \eta}{\partial \mu}\right)(Y - \mu)\right] \\
&= [X^T W X]^{-1} X^T W s
\end{aligned} \qquad (2.7).$$

Di mana hubungan antara W dan A sebagai berikut:

$$A = W \left(\frac{\partial \eta}{\partial \mu}\right) \text{ dan } \left(\frac{\partial \eta}{\partial \mu}\right) = diag \left(\frac{\partial \eta_i}{\partial \mu_i}\right)$$

maka didapat koefisien regresi dalam model regresi Poisson diestimasi menggunakan *Maximum Likelihood Estimation* (MLE)

$$\hat{\beta}_{MLE} = \left(X^t \widehat{W} X\right)^{-1} X^t \widehat{W} s \tag{2.8}$$

dengan:

$\hat{\beta}_{MLE}$: estimasi parameter β *Maximum Likelihood Estimation* (MLE),

X : variabel independen,

$\widehat{W}$: matriks diagonal, diag $[\mu_i]$,

s : $log(\hat{\mu}_i) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$.

2.3 Multikolinearitas

Dalam regresi Poisson, multikolinearitas di antara variabel independen merupakan asumsi yang perlu dipenuhi. Istilah multikolinearitas pertama kali diperkenalkan oleh Ragnar Frisch pada tahun 1924. Multikolinearitas dalam regresi Poisson terjadi ketika terdapat korelasi linier yang kuat antar variabel independen, sehingga dapat menyebabkan estimasi parameter menjadi tidak stabil dan analisis kurang akurat. Multikolinearitas dapat dideteksi dengan cara melihat nilai *Variance Inflation Factor* (VIF). Jika nilai VIF lebih dari 10, menunjukkan adanya masalah multikolinearitas. Rumus berikut dapat digunakan untuk mencari nilai VIF:

$$VIF = \frac{1}{1 - R_j^2} \tag{2.9}$$

R_j^2 adalah koefisien determinasi yang mengukur seberapa baik variabel independen X_j dapat menjelaskan variasi dalam variabel dependen ketika diregresikan terhadap variabel independen lainnya.

2.4 *Poisson James-Stein Estimator*

Poisson James-Stein Estimator (PJSE) adalah metode estimasi yang diusulkan untuk mengatasi masalah multikolinearitas dalam model regresi Poisson. Estimasi parameter menggunakan metode *Maximum Likelihood Estimation* (MLE) sering kali tidak memuaskan, terutama ketika terdapat multikolinearitas yang tinggi di antara variabel independen.

PJSE dikembangkan sebagai solusi dengan memanfaatkan konsep *shrinkage estimator*. *Estimator* ini dirancang untuk mengurangi inflasi varians yang dihasilkan oleh MLE dalam situasi di mana variabel penjelas saling berkorelasi tinggi. PJSE didefinisikan sebagai berikut:

$$\hat{\beta}_{PJSE} = c\hat{\beta}_{MLE} \tag{2.10),}$$

di mana (0 < c < 1) adalah faktor pengali yang ditentukan untuk mengurangi estimasi MLE yang didefinisikan sebagai berikut:

$$c = \frac{\left(\hat{\beta}'_{MLE}\hat{\beta}_{MLE}\right)}{\left(\hat{\beta}'_{MLE}\hat{\beta}_{MLE} + trace(S)^{-1}\right)} \tag{2.11}$$

dengan:

$\hat{\beta}_{MLE}$: estimasi MLE dari koefisien regresi,

S : $X^t\widehat{W}X$.

Bias $\hat{\beta}_{PJSE}$ dijelaskan sebagai berikut:

$$Bias\left(\hat{\beta}_{PJSE}\right) = E\left(\hat{\beta}_{PJSE}\right) - \beta \tag{2.12).}$$

Menggunakan persamaan 2.10 dan 2.11 ke dalam persamaan 2.12 didapatkan:

$$Bias(\hat{\beta}_{PJSE}) = E\left(\frac{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE})}{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE}+trace(S)^{-1})}\hat{\beta}_{MLE}\right) - \beta \qquad (2.13).$$

Secara sederhana, diperoleh persamaan:

$$Bias(\hat{\beta}_{PJSE}) = \left(-\frac{trace(S)^{-1}}{\beta'\beta+trace(S)^{-1}}\right)\beta \qquad (2.14).$$

Mean square error (MSE) dari $\hat{\beta}_{PJSE}$ dapat didefinisikan sebagai:

$$\begin{aligned} MSE(\hat{\beta}_{PJSE}) &= Var(\hat{\beta}_{PJSE}) + Bias(\hat{\beta}_{PJSE})'\, Bias(\hat{\beta}_{PJSE}) \\ &= c'Var(\hat{\beta}_{MLE})c + Bias(\hat{\beta}_{PJSE})'\, Bias(\hat{\beta}_{PJSE}) \qquad (2.15). \end{aligned}$$

Persamaan 2.11 dan 2.14 disubstitusikan ke persamaan 2.15, dapat diperoleh nilai MSE dari *James Stein Estimator* yaitu sebagai berikut:

$$\begin{aligned} MSE(\hat{\beta}_{PJSE}) = &\left(\frac{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE})}{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE}+trace(S)^{-1})}\right)'(trace(S)^{-1})\left(\frac{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE})}{(\hat{\beta}'_{MLE}\hat{\beta}_{MLE}+trace(S)^{-1})}\right) \\ &+\left(-\frac{trace(S)^{-1}}{\beta'\beta+trace(S)^{-1}}\beta\right)'\left(-\frac{trace(S)^{-1}}{\beta'\beta+trace(S)^{-1}}\beta\right) \qquad (2.16) \end{aligned}$$

disederhanakan, dengan skalar $MSE(\hat{\beta}_{PJSE})$ dihitung sebagai:

$$MSE(\hat{\beta}_{PJSE}) = \sum_{j=1}^{p}\frac{\alpha_j^4\lambda_j}{(\alpha_j^2\lambda_j+1)^2} + \sum_{j=1}^{p}\frac{\alpha_j^2}{\alpha_j^2{\lambda_j}^2+1} \qquad (2.17).$$

2.5 *Poisson Ridge Regression Estimator*

Poisson Ridge Regression (PRR) adalah sebuah pendekatan yang dimodifikasi untuk menangani masalah multikolinieritas atau korelasi tinggi di antara variabel independen, dengan menerapkan metode *ridge* dalam regresi Poisson. Metode *Ridge Regression* ini pertama kali diperkenalkan oleh Hoerl & Kennard pada tahun 1970, dan dimodifikasi oleh Mansson & Shukur pada tahun 2011.

Menurut Mansson & Shukur (2011), dalam penerapan metode *Poisson Ridge Regression*, prinsip yang digunakan adalah bahwa *Maximum Likelihood* memperkirakan jumlah kuadrat sisaan yang minimal, atau *Weight Sum of Square Error* (WSSE). *Estimator* $\boldsymbol{\hat{\beta}_{ML}}$ yang dihasilkan dianggap sebagai *estimator* yang optimal dalam konteks WSSE. Misalkan kita akan memilih *estimator* sembarang $\boldsymbol{\hat{B}}$ selain $\boldsymbol{\hat{\beta}_{ML}}$, dengan $\boldsymbol{\hat{B}}$ ini merupakan suatu vektor dari β maka dapat ditulis WSSE dari *estimator* ini adalah sebagai berikut:

$$\begin{aligned} \boldsymbol{\phi} &= \left(\boldsymbol{Y} - \boldsymbol{\hat{B}}\right)^{\boldsymbol{T}}\left(\boldsymbol{Y} - \boldsymbol{\hat{B}}\right) \\ &= \left(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\hat{\beta}_{ML}}\right)^{\boldsymbol{T}}\left(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\hat{\beta}_{ML}}\right) + \left(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}}\right)^{\boldsymbol{T}}\boldsymbol{X^T W \hat{X}}(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}}) \\ &= \boldsymbol{\phi_{min}} + \boldsymbol{\phi}\left(\boldsymbol{\hat{B}}\right) \qquad (2.18). \end{aligned}$$

Berdasarkan persamaan 2.18, $\boldsymbol{\phi}\left(\boldsymbol{\hat{B}}\right)$ adalah penambahan WSSE ketika *estimator* $\boldsymbol{\hat{\beta}_{ML}}$ digantikan dengan estimator $\boldsymbol{\hat{B}}$. Menurut Hoerl & Kennard (1970), untuk memperoleh nilai estimasi parameter dalam *Poisson Ridge Regression*, proses dilakukan dengan meminimalkan fungsi Lagrange ($\boldsymbol{F}$) menggunakan $\boldsymbol{\hat{B}^T\hat{B}}$, di mana $\boldsymbol{\phi}\left(\boldsymbol{\hat{B}}\right) = \boldsymbol{\phi 0}$, yang dapat dinyatakan dengan persamaan berikut:

$$\begin{aligned} &\boldsymbol{\phi}\left(\boldsymbol{\hat{B}}\right) = \boldsymbol{\phi_0} \\ &\left(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}}\right)^{\boldsymbol{T}}\boldsymbol{X^T W \hat{X}}\left(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}}\right) = \boldsymbol{\phi_0} \qquad (2.19). \end{aligned}$$

Selanjutnya, hal tersebut diformulasikan ke dalam bentuk Lagrangian, seperti yang ditunjukkan pada persamaan berikut:

$$\text{Min } \boldsymbol{F} = \boldsymbol{\hat{B}^T\hat{B}} + \left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)\left(\left(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}}\right)^{\boldsymbol{T}}\boldsymbol{X^T}\boldsymbol{\widehat{WX}}\left(\boldsymbol{\hat{B}} - \boldsymbol{\hat{\beta}_{ML}} - \boldsymbol{\phi_0}\right)\right)$$

$$= \widehat{\boldsymbol{B}}^T\widehat{\boldsymbol{B}} + \left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)\left(\widehat{\boldsymbol{B}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}}^T - \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}}^T\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T + \widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T - \boldsymbol{\phi_0}\right)$$

$$= \widehat{\boldsymbol{B}}^T\widehat{\boldsymbol{B}} + \left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)\left(\widehat{\boldsymbol{B}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} - \boldsymbol{2}\widehat{\boldsymbol{B}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T + \widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T - \boldsymbol{\phi_0}\right) \quad (2.20)$$

di mana:

$\left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)$: *multiple ragrange*

$\boldsymbol{\phi_0}$: konstanta non-negatif.

Setelah itu, persamaan 2.20 diturunkan terhadap $\widehat{\boldsymbol{B}}$ yang hasilnya disamadenganKan dengan nol

$$\frac{\boldsymbol{\partial F}}{\boldsymbol{\partial \widehat{B}}} = \frac{\boldsymbol{\partial}}{\boldsymbol{\partial \widehat{B}}}\left[\widehat{\boldsymbol{B}}^T\widehat{\boldsymbol{B}} + \left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)\left(\widehat{\boldsymbol{B}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} - \boldsymbol{2}\widehat{\boldsymbol{B}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T + \widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}^T - \boldsymbol{\phi_0}\right)\right]$$

$$= \boldsymbol{2}\widehat{\boldsymbol{B}} + \left(\frac{\boldsymbol{1}}{\boldsymbol{k}}\right)\left(\boldsymbol{2}\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} - \boldsymbol{2}\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}\right) = \boldsymbol{0} \quad (2.21).$$

$\widehat{\boldsymbol{B}}$ pada persamaan tersebut dikatakan sebagai *estimator* $\widehat{\boldsymbol{\beta}}_{\boldsymbol{PRRE}}$, maka persamaan *estimator* metode *Poisson Ridge Regression* dapat ditulis sebagai berikut:

$$\boldsymbol{2k}\widehat{\boldsymbol{B}} + \boldsymbol{2}\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} - \boldsymbol{2}\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}} = \boldsymbol{0}$$

$$\boldsymbol{2}\left(\boldsymbol{k}\widehat{\boldsymbol{B}} + \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}}\right) = \boldsymbol{2}\left(\boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}\right)$$

$$\boldsymbol{k}\widehat{\boldsymbol{B}} + \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{B}} = \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}$$

$$\widehat{\boldsymbol{B}}\left(\boldsymbol{kI} + \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\right) = \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}}$$

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{PRRE}} = \left(\boldsymbol{kI} + \boldsymbol{X}^T\widehat{\boldsymbol{W}}\boldsymbol{X}\right)^{-\boldsymbol{1}}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ML}} \quad (2.22)$$

di mana:

k : Parameter bias atau faktor *shrinkage*, $k > 0$,

I : Matriks identitas $p \times p$,

$\widehat{W}$: matriks diagonal, diag $\{\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \dots, \hat{\mu}_i\}$,

$\hat{\beta}_{ML}$: estimasi MLE dari koefisien regresi,

$\hat{\beta}_{PRRE}$: estimasi PRRE dari koefisien regresi.

Karakteristik dari *Poisson Ridge Regression Estimator* mencakup:

$$Bias(\hat{\beta}_{PRRE}) = (kI + X^T\widehat{W}X)^{-1}X^TWX\beta - \beta$$
$$= \left((kI + X^T\widehat{W}X)^{-1}X^TWX - I\right)\beta \qquad (2.23)$$

$$Bias(\hat{\beta}_{PRRE})^T Bias(\hat{\beta}_{PRRE}) = \beta^T \left((X^T\widehat{W}X + kI)^{-1}X^TWX - I\right)^T ((kI + X^T\widehat{W}X)^{-1}$$
$$X^TWX - I)\beta$$
$$= \beta^T \left(k(X^T\widehat{W}X + kI)^{-1}\right)^T \left(k(X^T\widehat{W}X + kI)^{-1}\right)\beta$$
$$= \beta^T k^2 (X^T\widehat{W}X + kI)^{-2}\beta \qquad (2.24),$$

dapat dituliskan dalam bentuk skalar sebagai berikut:

$$Bias(\hat{\beta}_{PRRE})^T Bias(\hat{\beta}_{PRRE}) = k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\lambda_j + k)^2} \qquad (2.25)$$

$$Var(\hat{\beta}_{PRRE}) = (kI + X^T\widehat{W}X)^{-2}(X^TWX)^2(X^TWX)^{-1}$$
$$= (kI + X^T\widehat{W}X)^{-2}X^TWX \qquad (2.26),$$

dapat ditulis dalam bentuk skalar berikut:

$$Var(\hat{\beta}_{PRRE}) = \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j + k)^2} \qquad (2.27).$$

Nilai MSE dari *Poisson Ridge Regression Estimator* yaitu sebagai berikut:

$$MSE(\hat{\beta}_{PRRE}) = Var(\hat{\beta}_{PRRE}) + Bias(\hat{\beta}_{PRRE})' \, Bias(\hat{\beta}_{PRRE})$$
$$MSE(\hat{\beta}_{PRRE}) = \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\lambda_j + k)^2} \qquad (2.28).$$

Penentuan parameter *Ridge k* diperlukan dalam model *estimator Poisson Ridge Regression* untuk mengatasi masalah multikolinearitas pada regresi Poisson. Beberapa metode telah diusulkan oleh penelitian terdahulu, sehingga dalam penelitian ini akan digunakan parameter *Ridge* sebagai berikut:

$$\hat{k}_1 = \frac{1}{\hat{\alpha}_{max}^2} \tag{2.29}$$

$$\hat{k}_2 = median(q_i) \tag{2.30}$$

di mana:

- $\hat{\alpha}_{max}^2$ merupakan nilai maksimum dari $\boldsymbol{\gamma\hat{\beta}_{ML}}$ dengan $\boldsymbol{\gamma}$ merupakan elemen vektor eigen dari matrik $\boldsymbol{X^t\widehat{W}X}$
- $q_i = \frac{\lambda max}{(n-p)\hat{\sigma}^2+\lambda_{max}\hat{\alpha}_i^2}$, dimana λmax merupakan nilai maksimum dari nilai eigen $\boldsymbol{X^t\widehat{W}X}$.
- $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i-\hat{\mu}_i)^2}{n-p-1}$ dimana p jumlah variabel *predictor*.

2.6 *Poisson Modified Kibria-Lukman Estimator*

Poisson Modified Kibria-Lukman Estimator (PMKLE) adalah pengembangan dari *estimator* KL yang digunakan untuk menangani multikolinearitas pada model regresi Poisson. Model regresi Poisson biasanya diaplikasikan pada data hitungan, di mana variabel dependen menunjukkan kejadian yang langka. PMKLE dibuat dengan mengganti komponen estimasi awal dari *estimator* KL menggunakan *estimator Ridge*. Estimasi parameter dalam PMKLE sebagai berikut:

$$\hat{\beta}_{PMKLE} = \left(X'\widehat{W}X + k\right)^{-1}\left(X'\widehat{W}X - k\right)\left(X'\widehat{W}X + k\right)^{-1}X'\widehat{W}X\hat{\beta}_{MLE} \quad (2.31).$$

Karakteristik dari *Poisson Modified Kibria-Lukman Estimator* mencakup:

$$E(\hat{\beta}_{PMKLE}) = \left(X'\widehat{W}X + k\right)^{-1}\left(X'\widehat{W}X - k\right)\left(X'\widehat{W}X + k\right)^{-1}X'\widehat{W}X\beta \tag{2.32}$$

$$\begin{aligned} Bias\left(\hat{\beta}_{PMKLE}\right) &= E\left(\hat{\beta}_{PMKLE}\right) - \beta \\ &= \left(X'\widehat{W}X + k\right)^{-1}\left(X'\widehat{W}X - k\right)\left(X'\widehat{W}X + k\right)^{-1}X'\widehat{W}X\beta - \beta \end{aligned}$$

$$= \left(X'\widehat{W}X + k\right)^{-2} k\left[-3X'\widehat{W}X - kI\right]\beta \qquad (2.33),$$

bias dapat dituliskan dalam bentuk skalar sebagai berikut:

$$Bias\left(\hat{\beta}_{PMKLE}\right) = k \sum_{j=1}^{p} \frac{\left(-3\lambda_j - k\right)\alpha_j}{\left(\lambda_j + k\right)^2} \qquad (2.34)$$

$$Var\left(\hat{\beta}_{PMKLE}\right) = \left(X'\widehat{W}X + k\right)^{-1}\left(X'\widehat{W}X - k\right)\left(X'\widehat{W}X + k\right)^{-1} X'\widehat{W}X\left(X'\widehat{W}X + k\right)^{-1}$$
$$\left(X'\widehat{W}X - k\right)\left(X'\widehat{W}X + k\right)^{-1}$$
$$= X'\widehat{W}X\left(X'\widehat{W}X - k\right)^2\left(X'\widehat{W}X + k\right)^{-4} \qquad (2.35),$$

dapat ditulis dalam bentuk skalar berikut:

$$Var\left(\hat{\beta}_{PMKLE}\right) = \sum_{j=1}^{p} \frac{\lambda_j\left(\lambda_j - k\right)^2}{\left(\lambda_j + k\right)^4} \qquad (2.36).$$

Efektivitas PMKLE dievaluasi menggunakan *Mean Squared Error* (MSE), yang menyatakan seberapa baik estimasi yang dihasilkan oleh *estimator* ini. MSE untuk PMKLE sebagai berikut:

$$MSE\left(\hat{\beta}_{PMKLE}\right) = Var\left(\hat{\beta}_{PMKLE}\right) + Bias\left(\hat{\beta}_{PMKLE}\right)' \, Bias\left(\hat{\beta}_{PMKLE}\right)$$
$$MSE(\hat{\beta}_{PMKLE}) = \sum_{j=1}^{p} \frac{\lambda_j\left(\lambda_j - k\right)^2}{\left(\lambda_j + k\right)^4} + k^2 \sum_{j=1}^{p} \frac{\left(3\lambda_j + k\right)^2 \alpha_j^2}{\left(\lambda_j + k\right)^4} \qquad (2.37).$$

Pemilihan nilai *k* biasanya dilakukan berdasarkan pendekatan yang meminimalkan *Mean Squared Error* (MSE) dari *estimator*. Beberapa metode yang digunakan untuk menentukan nilai *k* yaitu:

$$k_1 = \frac{1}{max\left(\alpha_j^2\right)} \qquad (2.38)$$

$$k_2 = \frac{p}{\sum\left(2\alpha_j^2 + \frac{1}{\lambda_j}\right)} \qquad (2.39)$$

$$k_3 = min\left(\frac{\lambda_i}{2\lambda_j \alpha_j^2 + 1}\right) \tag{2.40}$$

dimana λ adalah nilai eigen dari $X^t\widehat{W}X$ dan α^2 merupakan nilai dari $\boldsymbol{\gamma\widehat{\beta}_{ML}}$ dengan $\boldsymbol{\gamma}$ merupakan elemen vektor eigen dari matrik $X^t\widehat{W}X$.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada Semester Ganjil tahun ajaran 2024/2025 di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data penelitian yang digunakan yaitu data simulasi yang mengandung multikolinieritas yang dihasilkan dengan menggunakan *software R*. Adapun jumlah sampel yang digunakan adalah n = 20, 40, 60 dan 80 dengan p = 6 variabel dengan korelasi parsial (ρ = 0.3 dan 0.99) dan korelasi penuh (ρ = 0.99) serta dilakukan 1000 pengulangan. Persamaan berikut ini digunakan untuk menghasilkan setiap variabel independen Xp dengan menggunakan simulasi Monte Carlo untuk mendapatkan data multikolinearitas pada 6 variabel independen

$$X_p = \sqrt{(1-\rho^2)}Z_{ij} + \rho Z_{i(p+1)} \tag{3.1}$$

dengan i = 1, 2,..., n; j = 1, 2,..., p; $Z_{ij} \sim N(0,1)$; ρ = 0.3 dan 0.99.

Selanjutnya, variabel respons y_i mengikuti distribusi Poisson dengan parameter μ_i. Untuk setiap observasi i, variabel respons y_i dihasilkan menggunakan distribusi Poisson sebagai berikut:

$$P(\mu_i) = exp\left(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}\right) \quad (3.2).$$

Untuk mengevaluasi performa *estimator* menggunakan MSE sebagai berikut:

$$MSE = \frac{\sum_{i=1}^{R}\left(\hat{\beta}_i - \beta\right)'\left(\hat{\beta}_i - \beta\right)}{R} \quad (3.2)$$

di mana $\hat{\beta}$ adalah estimator β yang diperoleh dari PJSE, PRRE, dan PMKLE, dan R adalah jumlah pengulangan.

3.3 Metode Penelitian

Penelitian ini dilaksanakan dengan metode studi pustaka, dengan mempelajari berbagai buku referensi, baik berupa karya ilmiah maupun jurnal. Untuk mendukung proses perhitungan dan mendapatkan hasil yang akurat, penelitian ini menggunakan *software R*. Adapun tahapan yang dilakukan adalah sebagai berikut:

1. Melakukan simulasi data pada model Poisson dengan n = 20, 40, 60 dan 80,
2. Menguji multikolinearitas dengan mengevaluasi nilai VIF pada data simulasi,
3. Menentukan koefisien parameter (β) melalui analisis regresi Poisson menggunakan metode *Maximum Likelihood Estimation* (MLE) pada data simulasi dengan 3 dan 6 variabel independen yang saling berkorelasi,
4. Melakukan pendugaan parameter PJSE sebagai berikut:
   - Menentukan nilai parameter *c* pada PJSE,
   - Menghitung nilai estimasi PJSE,
   - Menghitung nilai MSE pada PJSE,
5. Melakukan pendugaan parameter PRRE sebagai berikut:
   - Menentukan nilai parameter *Ridge k* pada PRRE,

- Menghitung nilai estimasi PRRE untuk setiap nilai k,
- Menghitung nilai MSE pada PRRE dari beberapa parameter *Ridge* k,

6. Melakukan pendugaan parameter PMKLE sebagai berikut:
    - Menentukan nilai parameter k pada PMKLE,
    - Menghitung nilai estimasi PMKLE untuk setiap nilai k,
    - Menghitung nilai MSE pada PMKLE dari beberapa parameter k,
7. Membandingkan nilai MSE pada ketiga *estimator* berdasarkan hasil analisis yang telah dilakukan untuk mendapatkan metode terbaik.

V. KESIMPULAN

Berdasarkan hasil yang diperoleh, berikut ini adalah kesimpulan mengenai penggunaan metode PJSE, PRRE, dan PMKLE dalam analisis regresi Poisson:

1. Pada simulasi data korelasi parsial untuk ukuran sampel n = 20, performa metode *Poisson Ridge Regression Estimator* menggunakan nilai parameter k_2 dalam mengatasi multikolinearitas lebih baik dibandingkan dengan metode *Poisson James-Stein Estimator* dan *Poisson Modified Kibria-Lukman Estimator* karena menghasilkan nilai MSE yang lebih kecil Sedangkan pada ukuran sampel n = 40, 60 dan 80, performa metode *Poisson Modified Kibria-Lukman Estimator* menggunakan nilai parameter k_2 dalam mengatasi multikolinearitas lebih baik dibandingkan dengan metode lainnya karena menghasilkan nilai MSE yang lebih kecil.

2. Pada simulasi data korelasi penuh untuk ukuran sampel n = 20 dan 40 performa metode *Poisson Ridge Regression Estimator* menggunakan nilai parameter k_2 dalam mengatasi multikolinearitas lebih baik dibandingkan dengan metode *Poisson James-Stein Estimator* dan *Poisson Modified Kibria-Lukman Estimator* karena menghasilkan nilai MSE yang lebih kecil Sedangkan pada ukuran sampel n = 60 dan 80, performa metode *Poisson Modified Kibria-Lukman Estimator* menggunakan nilai parameter k_2 dalam mengatasi multikolinearitas lebih baik dibandingkan dengan metode lainnya karena menghasilkan nilai MSE yang lebih kecil.

DAFTAR PUSTAKA

Abonazel, M. R. 2023. New modified two-parameter Liu estimator for the Conway—Maxwell Poisson regression model. *Journal of Statistical Computation and Simulation.* **93**(12): 1976-1996.

Aladeitan, B. B., Adebimpe, O., Lukman, A. F., Oludoun, O., & Abiodun, O.E. 2021. Modified Kibria-Lukman (MKL) Estimator for the Poisson Regression Model. *F1000Research.* **10**(548): 1-19.

Amin, M., Akram, M. N., & Amanullah, M. 2020 . On the James-Stein Estimator for the Poisson Regression Model. *Communications in Statistics: Simulation and Computation.* **51**(10): 5596-5608.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T. C. 1985. *The Theory and Practice of Econometrics*. 2^{nd} Edition. John Wiley & Sons. USA.

Khoshsirat, S., & Kambhamettu, C. 2023 . Improving Normalization with the James-Stein Estimator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **42**(12): 2980-2992.

Lavery, M. R., Acharya, P., Sivo, S. A. & Xu, L. 2019. Number of predictors and Multicollinearity: What are Their Effects on Error and Bias in Regression?,.*Communications in Statistics-Simulation and Computation*. **48**(1): 27-38.

Lukman, A. F., Allohibi, J., Jegede, S. L., Adewuyi, E. T., Oke, S., & Alharbi, A. A. 2023. Kibria–Lukman-Type Estimator for Regularization and Variable Selection with Application to Cancer Data. *Mathematics.* **11**(23): 4795.

Månsson, K., & Shukur, G. 2011. A Poisson Ridge Regression Estimator. *Econimic Modelling*. **28**(4): 1475–1481.

Montgomery, D. C., Peck, E. A., & Vining, G. G. 2012. *Introduction to Linear Regression Analysis*. 4th Edition. John Willey and Sons, New York.

Myers, R. H. 1990. *Classical and Modern Regression with Applications*. 2nd Edition. Duxbury Press. California.

Oghenekevwe, E. H., Florence, A. K., & Abidemi, A. K., 2021. Poisson Ridge Regression Estimators: A Performance Test. *American Journal of Theoretical and Applied Statistics*. **10**(2): 111-121.

Pudjoatmodjo, B. Hendayun, M. 2016. Kehandalan Software Berdasarkan Data Sekunder Menggunakan Distribusi Poisson dan Kualifikasi Cronbach's Alpha. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi* (JNTETI). **5**(2): 53-62.

Sami, F., Amin, M., & Aljeddani, S. M. A., 2024 . Stein Estimation in the Conway-Maxwell Poisson Model with Correlated Regressors. *International Journal of Advanced and Applied Sciences*. **11**(7): 49- 56.

Wardah, T. L., Lubis, R. S., & Aprilia, R. 2022. Analisis Poisson Ridge Regression (PRR) pada Faktor yang Mempengaruhi Kecelakaan Lalu Lintas di Sumatera Utara. *Jurnal Matematika dan Pendidikan Matematika*. **5**(2): 154-160.