PEMODELAN REGRESI BINOMIAL NEGATIF MENGGUNAKAN ESTIMATOR *JACKKNIFE NEGATIVE BINOMIAL RIDGE REGRESSION* PADA DATA ANGKA KEMATIAN BAYI PROVINSI SULAWESI SELATAN

SKRIPSI

KEZIA AGRA PALINOAN
H051181504

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
APRIL 2023

PEMODELAN REGRESI BINOMIAL NEGATIF MENGGUNAKAN ESTIMATOR *JACKKNIFE NEGATIVE BINOMIAL RIDGE REGRESSION* PADA DATA ANGKA KEMATIAN BAYI PROVINSI SULAWESI SELATAN

SKRIPSI

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

KEZIA AGRA PALINOAN
H051181504

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

APRIL 2023

LEMBAR PERNYATAAN KEOTENTIKAN

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

Pemodelan Regresi Binomial Negatif Menggunakan Estimator *Jackknife Negative Binomial Ridge Regression* Pada Data Angka Kematian Bayi Provinsi Sulawesi Selatan

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

Makassar, 11 April 2023

Kezia Agra Palinoan

NIM. H051181504

PEMODELAN REGRESI BINOMIAL NEGATIF MENGGUNAKAN ESTIMATOR *JACKKNIFE NEGATIVE BINOMIAL RIDGE REGRESSION* PADA DATA ANGKA KEMATIAN BAYI PROVINSI SULAWESI SELATAN

Disetujui Oleh:

Pembimbing Utama	**Pembimbing Pertama**
Andi Kresna Jaya, S.Si., M.Si.	**Dr. Anna Islamiyati, S.Si., M.Si.**
NIP. 19731228 200003 1 001	**NIP. 19770808 200501 200 2**

Ketua Program Studi

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

Pada 11 April 2023

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	: Kezia Agra Palinoan
NIM	: H051181504
Program Studi	: Statistika
Judul Skripsi	: Pemodelan Regresi Binomial Negatif Menggunakan Estimator *Jackknife Negative Binomial Ridge Regression* Pada Data Angka Kematian Bayi Provinsi Sulawesi Selatan

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

DEWAN PENGUJI

1. Ketua : Andi Kresna Jaya, S.Si., M.Si. (.....)
2. Sekretaris : Dr. Anna Islamiyati, S.Si., M.Si. (.....)
3. Anggota : Dr. Nirwan, M.Si. (.....)
4. Anggota : Dra. Nasrah Sirajang, M.Si. (.....)

Ditetapkan di : Makassar
Tanggal : 11 April 2023

KATA PENGANTAR

Puji dan syukur selalu penulis panjatkan ke hadirat Tuhan Yang Maha Esa yang senantiasa menyertai dan melindungi penulis sepanjang penyusunan skripsi ini. Atas pertolongan-Nya yang yang selalu menyertai dan menuntun penulis hingga penulis dapat menyelesaikan skripsi ini yang berjudul **"Pemodelan Regresi Binomial Negatif Menggunakan Estimator *Jackknife Negative Binomial Ridge Regression* Pada Data Angka Kematian Bayi Provinsi Sulawesi Selatan"** sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis mengucapkan terima kasih sekaligus beryukur untuk kedua orang tua terkasih **Marthen Bake' Palinoan** dan **Debora Palin** yang selalu mendoakan, mendukung, memberikan cinta kasihnya dengan tulus ikhlas sampai saat ini. Ucapan terima kasih juga penulis sampaikan kepada seluruh keluarga besar dimana pun berada yang telah mendoakan.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga disampaikan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin berserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Bapak Andi Kresna Jaya, S.Si., M.Si.**, selaku Pembimbing Utama dan Pembimbing Pertama yaitu **Ibu Dr. Anna Islamiyati, S.Si., M.Si.**, yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya di tengah berbagai kesibukan dan prioritasnya untuk senantiasa memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.

5. **Bapak Dr. Nirwan, M.Si.** dan **Ibu Dra. Nasrah Sirajang, M.Si.**, selaku Tim Penguji yang telah memberikan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini serta waktu yang telah diberikan kepada penulis.
6. **Bapak Dr. Nirwan, M.Si.**, selaku Penasehat Akademik penulis dan sekaligus menjadi pembimbing utama skripsi penulis. Terima kasih atas kesabaran, kritik, serta motivasi yang selalu diberikan kepada Penulis selama menjalani pendidikan di Departemen Statistika.
7. Teman seperjuangan **Nopa, Restu,** dan **Manto** yang telah mendengarkan segala keluh kesah, memberikan nasehat, dukungan, bantuan, kerja sama, kenangan suka duka selama perkuliahan.
8. Teman-teman **Koperasi, Rael, Dian, Rifka, Yuyu, Hafis, Ika, Acca, Emi, Fani,** dan **Haksar** terima kasih atas segala kebersamaan selama perkuliahan. Terima kasih pula atas kerja sama, kenangan suka duka serta dukungan selama proses penyusunan skripsi ini.
9. Teman-teman **Statistika 2018** atas segala kebersamaan selama masa perkuliahan, kerja sama, kenangan suka duka serta dukungan selama proses perkuliahan hingga skripsi ini dapat selesai.
10. Teman-teman **GMKI Komisariat FMIPA Unhas** yang telah menjadi wadah pengembangan kualitas diri dan pelayanan bagi penulis. Terkhusus pengurus yang mendoakan dalam keberlangsungan pendidikan di kampus.
11. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu-persatu yang turut mendoakan, memberi dukungan moril maupun materil kepada penulis selama proses penyelesaian tugas akhir ini.
12. Dan tak lupa juga, terima kasih kepada **Kezia Agra Palinoan** yang selalu berjuang dan masih bertahan untuk mengerjakan skripsi ini. Terima kasih karena selalu berusaha memberikan yang terbaik.

Penulis menyadari bahwa tugas akhir ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang saran yang membangun diperlukan dalam penulisan selanjutnya. Akhirnya, penulis berharap skripsi ini dapat bermanfaat dalam pengembangan wawasan bidang ilmu statistika dan menjadi berkat untuk penulis dan pembaca.

Makassar, 11 April 2023

Kezia Agra Palinoan

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: Kezia Agra Palinoan
NIM	: H051181504
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

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

"Pemodelan Regresi Binomial Negatif Menggunakan Estimator *Jackknife Negative Binomial Ridge Regression* Pada Data Angka Kematian Bayi Provinsi Sulawesi Selatan"

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

Demikian pernyataan ini saya buat dengan sebenarnya.
Dibuat di Makassar pada tanggal 11 April 2023

Yang menyatakan

(Kezia Agra Palinoan)

ABSTRAK

Analisis regresi Negative Binomial adalah metode yang digunakan untuk menganalisis hubungan antara variabel prediktor terhadap variabel respon yang berdistribusi Poisson. Namun, regresi Poisson tidak dapat digunakan untuk memodelkan data dengan overdispersi maupun terdapat multikolinearitas. Untuk menyelesaikan masalah tersebut digunakan regresi Binomial Negatif dengan estimator *Jackknife Negative Binomial Ridge Regression*. Dalam penelitian ini, estimasi parameter regresi Binomial Negatif dengan estimator *Jackknife Negative Binomial Ridge Regression* diterapkan pada data tingkat kematian bayi di Sulawesi Selatan tahun 2017. Metode *Jackknife* berperan untuk mereduksi bias sehingga dapat diperoleh penaksiran parameter dengan bias yang kecil sedangkan metode *ridge* untuk menangani multikolinearitas. Metode pemilihan parameter *ridge* menggunakan nilai MSE terkecil. Model terbaik terbentuk pada model dengan parameter *ridge* 0.0081. Berdasarkan estimasi parameter yang terbentuk menunjukkan bahwa variabel jumlah bayi dengan berat badan lahir rendah (X_1), jumah bayi yang diberi ASI eksklusif (X_2), jumlah bayi yang mendapatkan vitamin A (X_3), jumlah cakupan pelayanan K4 pada ibu hamil (X_4), jumlah ibu hamil yang menerima imunisasi TT2+ (X_5), dan jumlah kelahiran (X_6) signifikan mempengaruhi jumlah kematian bayi.

Kata Kunci: *Jackknife Negative Binomial Ridge Regression*, Jumlah Kematian Bayi, Multikolinearitas, Regresi Binomial Negatif.

ABSTRACT

Poisson regression analysis is a method used to analyze the relationship between the predictor variable and the response variable with a Poisson distribution. However, Poisson regression cannot be used to model data with overdispersion or multicollinearity. To solve this problem, Negative Binomial regression is used with the Jackknife Negative Binomial Ridge Regression estimator. In this study, the estimation of the Negative Binomial regression parameter with the Jackknife Negative Binomial Ridge Regression estimator was applied to data on infant mortality rates in South Sulawesi in 2017. The Jackknife method plays a role in reducing bias so that parameter estimates with small bias can be obtained, while the ridge method handles multicollinearity. The ridge parameter selection method uses the smallest MSE value. The best model is formed in the model with a ridge parameter of 0.0081. Based on the estimated parameters formed, it shows that the variables are the number of babies with low birth weight (X_1), the number of babies who are exclusively breastfed (X_2), the number of babies who get vitamin A (X_3), the number of K4 service coverage for pregnant women (X_4), the number of pregnant women who received TT2+ immunization (X_5), and number of birth (X_6) is significantly affected the number of infant mortality.

***Keywords**: Jackknife Negative Binomial Ridge Regression, Number of Infant Mortality, Multicollinearity, Negative Binomial Regression.*

DAFTAR ISI

HALAMAN SAMPUL ..... **i**
HALAMAN JUDUL ..... **ii**
LEMBAR PERNYATAAN KEOTENTIKAN ..... **iii**
HALAMAN PERSETUJUAN PEMBIMBING ..... **iv**
HALAMAN PENGESAHAN ..... **v**
KATA PENGANTAR ..... **vi**
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI ..... **ix**
ABSTRAK ..... **x**
ABSTRACT ..... **xi**
DAFTAR ISI ..... **xii**
DAFTAR TABEL ..... **xiv**
DAFTAR LAMPIRAN ..... **xv**
BAB I PENDAHULUAN **1**
1.1 Latar Balakang 1
1.2 Rumusan Masalah 3
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
BAB II TINJAUAN PUSTAKA **5**
2.1 Regresi Binomial Negatif 5
2.2.1 Uji Signifikansi Parameter secara Parsial 6
2.2.2 Uji Signifikansi Parameter secara Serentak 6
2.2 Transformasi Data 7
2.3 Estimator *Negative Binomial Ridge Regression* 7
2.4 Estimator *Jackknife Negative Binomial Ridge Regression* 8
2.5 Parameter *Ridge* 9
2.6 Multikolinearitas 10
2.7 Angka Kematian Bayi 10
BAB III METODE PENELITIAN **12**
3.1 Jenis dan Sumber Data 12

3.2 Variabel Penelitian 12
3.3 Langkah Analisis 13
BAB IV HASIL DAN PEMBAHASAN 15
4.1 Analisis Data 15
4.1.1 Deskriptif Data 15
4.1.2 Uji Distribusi Poisson 16
4.2 Model Regresi Binomial Negatif 16
4.2.1 Estimasi Parameter Model 16
4.2.2 Pengujian Estimasi Parameter Model 20
4.2.3 Pengujian Signifikansi Parameter Model 21
4.3 Pengujian Multikolinearitas 22
4.4 Analisis *Jackknife Negative Binomial Ridge Regression* 22
4.4.1 Estimtor *Jackknife Negative Binomial Ridge Regression* 22
4.4.2 Estimasi Parameter Model Data Transformasi 24
4.4.3 Estimasi Parameter *Ridge k* 26
4.4.4 Estimasi Parameter *Jackknife Negative Binomial Ridge Regression* 28
4.4.5 Model Regresi Binomial Negatif Baru 29
BAB V PENUTUP 32
5.1 Kesimpulan 32
5.2 Saran 33
DAFTAR PUSTAKA 34
LAMPIRAN 37

DAFTAR TABEL

Tabel 4.1 Deskriptif Data 15
Tabel 4.2 Estimasi Model Regresi Binomial Negatif 20
Tabel 4.3 Nilai Uji Wald 21
Tabel 4.4 Nilai VIF 22
Tabel 4.5 Estimasi Model Regresi Binomial Negatif Data yang Ditransformasi 24
Tabel 4.6 Penduga Parameter γ_{ML} 25
Tabel 4.7 Penduga Parameter k 28
Tabel 4.8 Penduga Parameter *Jackknife Poisson Ridge Regression* 28
Tabel 4.9 Penduga Parameter β_j 30

DAFTAR LAMPIRAN

Lampiran 1. Data Kematian Bayi Provinsi Sulawesi Selatan Tahun 2017 38
Lampiran 2. Data Hasil Transformasi 39
Lampiran 3. Matriks **Z** 40
Lampiran 4. *Syntax R Jackknife Poisson Ridge Regression* 41

BAB I
PENDAHULUAN

1.1 Latar Balakang

Tingkat kematian adalah prediktor konsekuensi kesehatan di masa depan. Angka kematian dapat menggambarkan seberapa baik sektor kesehatan suatu masyarakat. Pada dasarnya, tampaknya ada penyebab kematian tidak langsung dan langsung. Namun dalam praktiknya, banyak faktor berinteraksi untuk menentukan tingkat kematian dalam populasi (Ismail dkk., 2019). Angka Kematian Bayi (AKB) adalah jumlah bayi yang meninggal sebelum berumur satu tahun pada tahun tertentu per 1000 kelahiran hidup. Karena bayi baru lahir adalah kelompok usia yang paling terpapar pengaruh perubahan lingkungan dan sosial ekonomi, ciri dari kondisi sosial ekonomi masyarakat lokal dapat dilihat dari angka kematian bayinya (Wardani dan Wulandari, 2020). Tujuan *Millenium Development Targets* (MDGs) adalah meminimalkan angka kematian, dengan tujuan menurunkan angka kematian bayi hingga dua pertiganya selama tahun 1990 dan 2015. Salah satu indikatornya adalah angka kematian bayi (BPS, 2016) dalam (Cabral dkk., 2019).

Angka kematian bayi merupakan ukuran keberhasilan suatu negara dalam menentukan status kesehatannya. Namun karena pencapaian penurunan angka kematian ibu negara dari industri maju, angka kematian bayi kini dianggap sebagai indikator unggulan untuk menilai kepuasan konsumen pelayanan kesehatan (Cabral dkk., 2019). Angka kematian bayi merupakan statistik yang berguna untuk menentukan status kesehatan anak, posisi ekonomi penduduk suatu wilayah, dan target meminimalkan kematian bayi. Mencari tahu penyebab kematian bayi merupakan satu dari banyak cara untuk mengatasi masalah ini dan menguranginya (Elyna, 2012) dalam (Ismail dkk., 2019).

Berdasarkan data Profil Kesehatan tahun 2020 Dinas Kesehatan Provinsi Sulawesi Selatan, di Provinsi Sulawesi Selatan sendiri tercatat jumlah angka kematian bayi pada tahun 2020 mencapai angka 754 bayi atau 4,87 per 1000 kelahiran hidup. Ini adalah angka yang signifikan. Banyak faktor yang mempengaruhi kematian bayi dan ibu, antara lain proporsi jumlah ibu bersalin yang mendapat pelayanan kesehatan nifas, persentase persalinan yang dilakukan oleh

tenaga kesehatan, proporsi ibu hamil yang mendapat tablet Fe3, proporsi ibu hamil peserta program K4, persentase penanganan komplikasi neonatal, persentase penanganan komplikasi kebidanan, dan persentase bayi yang mendapat ASI eksklusif (Fitryanti dan Kurniawan, 2019) dalam (Islami dkk., 2021).

Sebagai subjek penelitian terapan, aplikasi statistika berkembang cukup pesat di era globalisasi. Analisis regresi adalah salah satu teknik statistik yang paling populer digunakan untuk memperoleh temuan penelitian. Hal ini disebabkan karena analisis regresi merupakan teknik statistik yang digunakan untuk menganalisis hubungan dari satu atau banyak variabel prediktor $X_1, X_2, \ldots, X_m$ dan variabel respon Y (Jus'at, 2019) dalam (Jao dkk., 2022). Analisis regresi digunakan di bidang kesehatan untuk mengetahui variabel-variabel model regresi yang memerangi kematian bayi baru lahir dan menyelidiki variabel yang mempengaruhi tingkat kematian bayi. Jumlah kematian bayi baru lahir di bawah satu tahun dapat digunakan untuk menghitung angka kematian bayi (Dinas Kesehatan Sulawesi Selatan, 2017) dalam (Jao dkk., 2022).

Karena jumlah kematian bayi baru lahir bersifat diskrit, maka terdapat dalam data yang berdistribusi Poisson. Salah satu model regresi yang dapat digunakan untuk mengeksplorasi hubungan yang melibatkan faktor prediktor dan variabel target secara data diskrit adalah Model Regresi Poisson (Aulele, 2012) dalam (Jao dkk., 2022). Untuk melakukan analisis regresi Poisson, beberapa asumsi harus dipenuhi, salah satunya adalah ekuidispersi. Hitung data sering menunjukkan varians yang lebih besar daripada rata-rata, suatu kondisi yang dikenal sebagai overdispersi (Cameron dan Trivedi, 1998) dalam (Islami dkk., 2021). Regresi Poisson tidak dapat digunakan untuk memodelkan data dengan overdispersi. Salah satu strategi untuk mengatasi overdispersi pada regresi Poisson adalah regresi Binomial Negatif Bivariat (Islami dkk., 2021).

Regresi Binomial Negatif merupakan pendekatan yang dapat menangani overdispersi pada variabel dependen model regresi Poisson. Model regresi Binomial Negatif dibangun berdasarkan campuran antara distribusi Poisson dan distribusi Gamma (Hilbe, 2014) dalam (Fathurahman, 2022).

Estimasi parameter regresi Poisson, di sisi lain, menggunakan Kemungkinan Maksimum (ML). Namun, prosedur ini sering menghasilkan temuan yang salah.

Multikolinearitas adalah salah satunya. Akibatnya, hasil estimasi parameter memiliki varians yang tidak nol. Karena data memiliki sedikit pengamatan dan banyak faktor yang mempengaruhinya, terdapat indikator kekhawatiran terjadi multikolinearitas dalam data kematian bayi. Pendekatan metode *ridge* yang dikembangkan oleh Hoerl dan Kennard merupakan salah satu metode untuk menangani multikolinearitas (Hoerl dan Kennard, 1970) dalam (Munawaroh, 2016). Kemudian metode *Jackknife* berperan untuk mereduksi bias sehingga dapat diperoleh penaksiran parameter dengan bias yang kecil.

Penelitian sebelumnya telah dilakukan oleh Fathurahman (2022) mengenai regresi Binomial Negatif untuk memodelkan kematian bayi di Kalimantan Timur, Wardani dan Wulandari (2020) mengenai pemodelan *Negative Binomial Regression* pada data jumlah kematian bayi di Kabupaten Jombang, Silvia (2021) mengenai analisis angka kelahiran bayi pada remaja yang berusia 15-19 tahun menggunakan regresi Binomial Negatif, dan Savitri, dkk. (2022) mengenai determinan jumlah kematian balita akibat pneumonia di Indonesia tahun 2019 dengan pendekatan *Generalized Poisson Regression*.

Berdasarkan uraian di atas, peneliti terdorong untuk melakukan penelitian yang berkaitan dengan judul penelitian ini "Pemodelan Regresi Binomial Negatif Menggunakan Estimator *Jackknife Negative Binomial Ridge Regression* pada Data Angka Kematian Bayi Provinsi Sulawesi Selatan". Penulis akan menggunakan model regresi Binomial Negatif untuk mengatasi multikolinearitas yang terjadi pada data angka kematian bayi dengan menggunakan estimatior *Jackknife Negative Binomial Ridge Regression*.

1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah dari penelitian ini adalah bagaimana model terbaik dari regresi Binomial Negatif yang menggunakan estimator *Jackknife Negative Binomial Ridge Regression* pada Angka Kematian Bayi di Provinsi Sulawesi Selatan tahun 2017?

1.3 Batasan Masalah

Adapun batasan masalah yang akan dibahan dalam penelitian ini yaitu:

1. Data yang digunakan yaitu kasus Angka Kematian Bayi yang berada di Provinsi Sulawesi Selatan yang terdiri dari 24 kabupaten/kota pada tahun 2017.
2. Parameter *ridge k* yang digunakan sebanyak 4 parameter.
3. Perhitungan menggunakan *software RStudio.*
4. Pemilihan model terbaik berdasakan nilai MSE terkecil.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah memperoleh model terbaik dari regresi Binomial Negatif yang menggunakan estimator *Jackknife Negative Binomial Ridge Regression* pada Angka Kematian Bayi di Provinsi Sulawesi Selatan tahun 2017.

1.5 Manfaat Penelitian

Adapun manfaat yang akan dicapai dari hasil penelitian ini, sebagai berikut:

1. Menambah ilmu dan wawasan terkait bentuk model regresi Binomial Negatif menggunakan estimatior *Jackknife Negative Binomial Ridge Regression* untuk mengatasi multikolinearitas pada data angka kematian bayi.
2. Menjadi pertimbangan pemerintah untuk mengambil kebijakan terkait angka kematian bayi khususnya di Provinsi Sulawesi Selatan.

BAB II
TINJAUAN PUSTAKA

2.1 Regresi Binomial Negatif

Menurut Keswari, dkk. (2014) Model regresi Binomial Negatif menggabungkan distribusi Poisson dan Gamma. Asumsikan *y* adalah variabel dependen dan memiliki distribusi Binomial Negatif. Regresi Binomial Negatif menggunakan model GLM bersama dengan tautan log dengan asumsi bentuk berikut dengan $\boldsymbol{y}$ dan $\boldsymbol{\beta}$ merupakan matriks vektor berukuran $n \times 1$, dan $\mathbf{X}$ adalah matriks berukuran $n \times p$:

$$\boldsymbol{\mu} = \exp \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta} \tag{2.1}$$

$$\boldsymbol{y} = \log \boldsymbol{\mu} = \boldsymbol{\eta} = \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}$$

$$\eta_i = \ln(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$$

dimana:

η_i = fungsi penghubung

μ_i = nilai tengah peubah respon ke-*i*

β_0 = konstanta

$\beta_1, \beta_2, \ldots, \beta_p$ = koefisien regresi variabel prediktor ke-1 sampai *p*

x_{ij} = nilai variabel prediktor ke-*j* amatan ke-*i*

Menurut Majore, dkk. (2020) Karena mean dan variasi model Binomial Negatif tidak harus sama, model regresi Binomial Negatif tampaknya lebih fleksibel daripada model Poisson. Berikut ini adalah fungsi peluang Binomial Negatif:

$$f(y; \mu, a) = \frac{\Gamma(y + a^{-1})}{\Gamma(a^{-1})y!}\left(\frac{a\mu}{1 + a\mu}\right)^{y}\left(\frac{1}{1 + a\mu}\right)^{a^{-1}}$$

Dengan *a* merupakan parmeter dispersi, $a > 0$. Jika $a \to 0$, maka distribusi ini mendekati Poisson (μ). Sebaran Binomial Negatif memiliki rataan $\mathrm{E(Y)} = \mu$ dan ragam $\mathrm{Var(Y)} = \mu + a\mu^2$.

2.2.1 Uji Signifikansi Parameter secara Parsial

Uji signifikansi parameter secara parsial adalah uji untuk masing-masing parameter yang berpengaruh terhadap model yang terbentuk pada regresi Binomial Negatif. Uji yang digunakan untuk melakukan uji signifikansi parameter secara parsial adalah uji *Wald*. Dalam uji ini digunakan hipotesis sebagai berikut.

Hipotesis:

H_0: $\beta_j = 0, j = 1, 2, \ldots, p$ (variabel prediktor ke-*j* tidak berpengaruh secara signifikan)

H_1: $\beta_j \neq 0, j = 1, 2, \ldots, p$ (variabel prediktor ke-*j* berpengaruh secara signifikan)

Statistik uji:

$$W_j = \left[\frac{\hat{\beta}_j}{SE\left(\hat{\beta}_j\right)}\right]^2 \tag{2.2}$$

dengan:

$\hat{\beta}_j$ = nilai estimasi parameter β_j

$SE\left(\hat{\beta}_j\right)$ = standar *error* dari β_j

Kriteria pengujian:

Tolak H_0 jika $W_j > \chi^2_{(\alpha,\, \text{db})}$ atau *p-value* < α.

2.2.2 Uji Signifikansi Parameter secara Serentak

Uji serentak terhadap signifikan parameter regresi Binomial Negatif dilakukan menggunakan statistik uji rasio *likelihood*. Dalam uji ini digunakan hipotesis sebagai berikut.

Hipotesis:

H_0: $\beta_0 = \beta_1 = \ldots = \beta_p = 0, j = 1, 2, \ldots, p$ (semua variabel prediktor tidak berpengaruh secara signifikan)

H_1: $\beta_j \neq 0, j = 1, 2, \ldots, p$ (terdapat minimal satu variabel prediktor yang berpengaruh secara signifikan)

Statistik uji:

$$LR = -2\left(L_{null} - L_{full}\right) \tag{2.3}$$

dengan:

L_{null} = Nilai estimasi dari fungsi log-*likelihood* setelah menstubtitusikan nilai taksiran parameter yang diperoleh dari taksiran intersep model saja

L_{full} = Nilai estimasi dari fungsi log-*likelihood* setelah mensubstitusikan nilai-nilai taksiran parameter yang diperoleh dari taksiran *full* model

Kriteria pengujian:

Tolak H_0 jika $LR \geq \chi^2_{(\alpha,\, \text{db})}$ atau *p-value* < α.

2.2 Transformasi Data

Modifikasi sederhana dari pembakuan atau standarisasi variabel ini adalah transformasi korelasi. Pemusatan merupakan perbedaan antara masing-masing pengamatan dengan rata-rata dari semua pengamatan untuk variabel. Sedangkan penskalaan meliputi gambaran pengamatan pada kesatuan (unit) standar deviasi dari pengamatan untuk variabel. Prosedur pemusatan atau centering dilakukan dengan menghilangkan β_0 yang membuat perhitungan untuk mencari model regresi menjadi lebih sederhana. Prosedur penskalaan dilakukan dengan cara mentransformasikan variabel independen X dalam bentuk berikut (Anggraini, dkk., 2019).

$$x_{ij}^{*} = \frac{x_{ij} - \bar{x}_j}{\sqrt{n-1}S_j} \tag{2.4}$$

2.3 Estimator *Negative Binomial Ridge Regression*

Dalam Singh, dkk. (1986) diberikan matriks **G** berukuran $p \times p$ dengan elemen vector eigen dari $\mathbf{X}^{\mathrm{T}}\widehat{\mathbf{W}}\mathbf{X}$ dan Λ_{NBRR} adalah matriks diagonal $(\lambda_{1\text{NBRR}}, \ldots, \lambda_{p\text{NBRR}})$ yang bersesuaian dengan matriks **G**, maka:

$$\begin{aligned}\Lambda_{\text{NBRR}} &= \mathbf{G}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\widehat{\mathbf{W}}\mathbf{X}\mathbf{G} \\ &= (\mathbf{X}\mathbf{G})^{\mathrm{T}}\widehat{\mathbf{W}}(\mathbf{X}\mathbf{G}) \\ &= \mathbf{Z}^{\mathrm{T}}\widehat{\mathbf{W}}\mathbf{Z}\end{aligned}$$

dengan $\mathbf{Z} = \mathbf{X}\mathbf{G}$

Kemudian dapat ditulis bentuk $\hat{\boldsymbol{\gamma}}_{ML}$ sebagai berikut.

$$\hat{\boldsymbol{\gamma}}_{ML} = \left(\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{Z}\right)^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \Lambda_{NBRR}^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$

dengan $\hat{\boldsymbol{s}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{ML}$

Sehingga estimator $\hat{\boldsymbol{\beta}}_{ML}$ dituliskan sebagai berikut.

$$\hat{\boldsymbol{\beta}}_{ML} = \left(\mathbf{X}^T\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{X}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \left(\mathbf{X}^T\right)^{-1}\left(\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{X}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \left[\Lambda_{NBRR}\left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\right]^{-1}\mathbf{X}^{-1}\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \Lambda_{NBRR}^{-1}\left[\left(\mathbf{Z}^T\mathbf{Z}\right)^{-1}\right]^{-1}\mathbf{X}^{-1}\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \Lambda_{NBRR}^{-1}\mathbf{Z}^T\mathbf{Z}\mathbf{X}^{-1}\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \Lambda_{NBRR}^{-1}\mathbf{Z}^T\mathbf{X}\mathbf{G}\mathbf{X}^{-1}\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \mathbf{G}\Lambda_{NBRR}^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \mathbf{G}\hat{\boldsymbol{\gamma}}_{ML} \qquad (2.5)$$

Estimator $\hat{\boldsymbol{\gamma}}_{NBRR}$ didapatkan dengan menambahkan konstanta bias $k\mathbf{I}$.

$$\hat{\boldsymbol{\gamma}}_{NBRR} = \left(\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{Z} + k\mathbf{I}\right)^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\hat{\boldsymbol{s}}$$
$$= \left(\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{Z} + k\mathbf{I}\right)^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}_{ML}$$
$$= \mathbf{B}^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{X}\mathbf{G}\hat{\boldsymbol{\gamma}}_{ML}$$
$$= \mathbf{B}^{-1}\mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{Z}\hat{\boldsymbol{\gamma}}_{ML}$$
$$= \mathbf{B}^{-1}(\mathbf{B} - k\mathbf{I})\hat{\boldsymbol{\gamma}}_{ML}$$
$$= \left(\mathbf{I} - k\mathbf{B}^{-1}\right)\hat{\boldsymbol{\gamma}}_{ML} \qquad (2.6)$$

dengan $\mathbf{B} = \mathbf{Z}^T\widehat{\mathbf{W}}\mathbf{Z} + k\mathbf{I}$

2.4 Estimator *Jackknife Negative Binomial Ridge Regression*

Metode *Jackknife* berperan dapat mereduksi bias sehingga dapat diperoleh penaksir parameter dengan bias yang kecil. Misalkan $\hat{\boldsymbol{s}}_{-i}$, $\mathbf{Z}_{-i}$, $\widehat{\mathbf{W}}_{-i}$, ditunjukkan masing-masing, vektor $\boldsymbol{s}$ dengan mengeluarkan baris ke-i, matriks $\mathbf{Z}$ dengan mengeluarkan baris ke-i, matriks $\mathbf{W}$ dengan mengeluarkan kedua baris dan kolom ke-i. Didefinisikan:

$$\hat{\boldsymbol{\gamma}}_{\text{NBRR-}i} = \left(\mathbf{Z}_{-i}^{\text{T}}\widehat{\mathbf{W}}_{-i}\mathbf{Z}_{-i} + \boldsymbol{k}\mathbf{I}\right)^{-1}\mathbf{Z}_{-1}^{\text{T}}\widehat{\mathbf{W}}_{-i}\hat{\mathbf{s}}_{-i} \tag{2.7}$$

Persamaan (2.5) merupakan estimator *Negative Binomial Ridge Regression* dengan mengeluarkan observasi ke-*i*, dimana:

$\mathbf{Z}_{-1}^{\text{T}}\widehat{\mathbf{W}}_{-i}\hat{\mathbf{s}}_{-i} = \mathbf{Z}^{\text{T}}\widehat{\mathbf{W}}\hat{\mathbf{s}} - \boldsymbol{z}_i^{T}\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{s}}_i$ dan $\left(\mathbf{Z}_{-i}^{\text{T}}\widehat{\mathbf{W}}_{-i}\mathbf{Z}_{-i} + \boldsymbol{k}\mathbf{I}\right)^{-1}$ merupakan teorema dari Sherman-Morrison Woodbury menggunakan invers jumlah matriks. Sehingga Persamaan (2.5) menjadi:

$$\hat{\boldsymbol{\gamma}}_{\text{NBRR-}i} = \hat{\boldsymbol{\gamma}}_{\text{NBRR}} - \frac{\left(\mathbf{Z}^{\text{T}}\widehat{\mathbf{W}}\mathbf{Z} + \boldsymbol{k}\mathbf{I}\right)^{-1}\boldsymbol{z}_i^{\text{T}}\hat{\boldsymbol{\mu}}_i\left(\hat{\boldsymbol{s}}_i - \boldsymbol{z}_i\hat{\boldsymbol{\gamma}}_{\text{NBRR}}\right)}{1 - \boldsymbol{z}_i^{\text{T}}\left(\mathbf{Z}^{\text{T}}\widehat{\mathbf{W}}\mathbf{Z} + \boldsymbol{k}\mathbf{I}\right)^{-1}\boldsymbol{z}_i\hat{\boldsymbol{\mu}}_i} \tag{2.8}$$

Estimator *Jackknifed Negative Binomial Ridge Regression* dapat dituliskan sebagai berikut.

$$\hat{\boldsymbol{\gamma}}_{\text{JNBR}} = \left(\mathbf{I} + \boldsymbol{k}^2\mathbf{B}^{-2}\right)\hat{\boldsymbol{\gamma}}_{\text{ML}} \tag{2.9}$$

Sehingga variansi untuk *Jackknifed Negative Binomial Ridge Regression* yaitu:

$$\text{Var}\left(\hat{\boldsymbol{\gamma}}_{\text{JNBR}}\right) = \left(\mathbf{I} + \boldsymbol{k}^2\mathbf{B}^{-2}\right)\Lambda_{\text{NBRR}}^{-1}\left(\mathbf{I} + \boldsymbol{k}^2\mathbf{B}^{-2}\right) \tag{2.10}$$

Estimator dari *Jackknifed Negative Binomial Ridge Regression* bersifat bias. Bias untuk *Jackknife Negative Binomial Ridge Regression* yaitu:

$$\text{Bias}\left(\hat{\boldsymbol{\gamma}}_{\text{JNBR}}\right) = -\boldsymbol{k}^2\mathbf{B}^{-2}\hat{\boldsymbol{\gamma}}_{\text{ML}} \tag{2.11}$$

Mean Square Error dari *Jackknife Negative Binomial Ridge Regression*:

$$\text{MSE}\left(\hat{\boldsymbol{\gamma}}_{\text{JNBR}}\right) = \left(\mathbf{I} + \boldsymbol{k}\mathbf{B}^{-1}\right)\Lambda_{\text{NBRR}}^{-1}\left(\mathbf{I} + \boldsymbol{k}\mathbf{B}^{-1}\right) + \boldsymbol{k}^4\mathbf{B}^{-2}\hat{\boldsymbol{\gamma}}_{\text{ML}}\hat{\boldsymbol{\gamma}}_{\text{ML}}^{\text{T}}\mathbf{B}^{-2} \tag{2.12}$$

2.5 Parameter *Ridge*

Penentuan nilai parameter *ridge* digunakan dalam penanganan multikolinearitas dalam regresi *ridge*. Berikut beberapa parameter *ridge*:

$$k_1 = \frac{1}{\hat{\alpha}_{max}^2} \tag{2.13}$$

$$k_2 = median\left(s_j\right) \tag{2.14}$$

$$k_3 = median\left(q_j\right) \tag{2.15}$$

$$k_4 = max\left(q_j\right) \tag{2.16}$$

dimana:

$\hat{\alpha}_{max}^2$ = nilai maksimum dari $\left(\mathbf{G}\boldsymbol{\beta}_{\text{ML}}^{*}\right)^2$

$$s_j = \frac{t_j \sigma^2}{(n - p)\sigma^2 + t_j \hat{\alpha}_j^2}$$

t_j = nilai eigen dari $\mathbf{X}^T\mathbf{X}$

$$\sigma^2 = \frac{(\boldsymbol{y} - \mu)^2}{(n - p - 1)}$$

$$q_j = \frac{\lambda_{max}}{(n - p)\sigma^2 + \lambda_{max} \hat{\alpha}_j^2}$$

λ_{max} = nilai eigen maksimum dari $\mathbf{G}$

2.6 Multikolinearitas

Menurut Munawaroh (2018) istilah multikolinieritas pertama kali diperkenalkan oleh Ragner Frisch adalah hubungan linier yang sangat tinggi antar variabel prediktor suatu model regresi. Karena koefisien regresi yang dihitung tidak benar-benar mantap dan variabel koefisien regresi tampaknya sangat besar karena multikolinearitas, pendekatan kuadrat terkecil tidak memadai. *Variance Inflation Factor* (VIF) atau nilai toleransi (TOL) dapat digunakan untuk mendeteksi multikolinearitas. Suatu dikatakan adanya kolinieritas tinggi jika memiliki nilai VIF > 10 atau nilai TOL < 0.1.

$$\text{VIF}_j = \frac{1}{1 - \text{R}_j^2}; j = 1, 2, \ldots, p \tag{2.17}$$

Dengan R adalah nilai koefisien determinasi antara variabel prediktor ke-*j* dan variabel prediktor lainnya. *j* adalah indeks untuk variabel prediktor dan *p* adalah banyaknya variabel prediktor.

2.7 Angka Kematian Bayi

Banyaknya bayi yang meninggal pada satu tahun disebut dengan angka kematian (mortalitas). Sebagai akibat dari suatu keadaan seperti penyakit atau sebab lain. Angka kematian di suatu wilayah dapat memberikan gambaran tentang evolusi status kesehatan masyarakat, yang sering digunakan untuk mengevaluasi keberhasilan inisiatif pengembangan pelayanan kesehatan. Kematian bayi didefinisikan sebagai kematian yang terjadi antara usia nol sampai satu tahun.

Kematian bayi dibagi dalam dua kategori: neonatal (terjadi dalam 28 hari pertama) dan postnatal (terjadi antara 28 dan 365 hari). Penyebab kematian neonatal endogen atau eksternal dibedakan. Kematian Endogen (Kematian Neonatal) adalah jenis kematian bayi baru lahir yang terjadi selama bulan pertama kehidupan dan juga terutama disebabkan oleh variabel yang diturunkan oleh bayi saat lahir dan diturunkan dari orang tuanya selama kehamilan. Kematian eksogen (Post Neonatal Death) adalah kematian bayi antara usia 1 bulan sampai 1 tahun.

Kematian bayi didefinisikan sebagai kematian yang terjadi antara waktu bayi lahir sampai dengan usia satu tahun (Amiruddin dan Hasmi, 2013). Kematian bayi diklasifikasikan sebagai neonatal (dalam 28 hari pertama) atau postneonatal (rentang 28 dan 365 hari) (Rachmawati, 2014). Secara umum penyebab kematian bayi ada dua jenis yaitu endogen dan eksogen. Kematian bayi endogen (kematian neonatus) adalah kematian hayi yang terjadi dalam bulan pertama kehidupan, dan penyebabnya seringkali bersifat variabel yang dibawa oleh bayi sejak lahir dan diturunkan dari orang tuanya selama kehamilan. Demikian juga, ada kematian bayi eksogen. Kematian Pasca Neonatal adalah kematian hayi yang terjadi setelah usia bayi baru lahir yaitu satu bulan sampai kurang dari satu tahun, dengan penyebabnya adalah penyebab lingkungan eksternal (Fathurahman, 2022).

Menurut Munawaroh (2018) Angka kematian bayi (AKB) adalah persentase kematian bayi di bawah usia satu tahun per 1000 kelahiran hidup pada tahun tertentu. Ada beberapa variabel yang berkontribusi terhadap kematian bayi. Ada dua jenis kematian bayi berdasarkan penyebabnya, yaitu endogen dan eksogen. Kematian endogen (neonatal) adalah kematian bayi yang terjadi dalam bulan pertama kehidupan dan biasanya disebabkan oleh faktor-faktor yang diturunkan melalui orang tuanya selama kehamilan. Kematian bayi eksogen (post neonatal) didefinisikan sebagai kematian bayi baru lahir antara usia satu bulan sampai satu tahun karena keadaan yang berhubungan dengan pengaruh lingkungan sekitarnya. Beberapa faktor yang mempengaruhi angka kematian bayi yaitu Bayi dengan Berat Lahir Rendah (BBLR), pemberian Asi Ekslusif. Air Susu Ibu (ASI) pada bayi, Pemberian vitamin A pada bayi, cakupan Ibu melakukan kunjungan bayi, pemberian imunisasi *Tetanus Toxoid* (TT) pada ibu minimal dua kali pada saat kehamilan.