

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perserikatan Bangsa-Bangsa (PBB) telah mencanangkan suatu program pembangunan berkelanjutan yang bernama *Sustainable Development Goals* atau lebih dikenal dengan SDGs (United Nations, 2020). SDGs merupakan komitmen global dan nasional yang bertujuan untuk meningkatkan kesejahteraan masyarakat dengan 17 tujuan dan 169 target yang harus dicapai per tahun 2030. Pada SDGs Tujuan ke-3, yaitu “Kehidupan Sehat dan Sejahtera”, terdapat 13 target yang harus dicapai untuk dapat menjamin kehidupan yang sehat dan meningkatkan kesejahteraan seluruh penduduk di semua usia. Salah satu target tersebut yang tertuang dalam Target 3.2 adalah menurunkan angka kematian bayi baru lahir hingga 12 per 1000 kelahiran hidup (SDGs Indonesia, 2023). Pencapaian tujuan dan target SDGs dipengaruhi oleh kualitas sumber daya manusia itu sendiri sehingga diperlukan perhatian yang khusus dalam sektor pendidikan, kesehatan, dan lingkungan (Sumargo dkk., 2018).

Bayi termasuk kelompok yang memerlukan perhatian khusus di sektor kesehatan. Kesehatan bayi berperan penting dalam pencapaian SDGs karena mereka sedang berada dalam fase *golden age*. Fase *golden age* atau masa keemasan ini merupakan fase penentu kehidupan sehingga diperlukan dukungan yang kuat untuk mengoptimalkan perkembangan bayi (Afnita & Latipah, 2021). Sejak lahir hingga bertambahnya usia, bayi mengalami pertumbuhan dan perkembangan yang sangat pesat, terutama bayi di tahun pertama (Jannah dkk., 2017). Pada usia ini, bayi mulai mengembangkan keterampilan motorik, seperti merangkak, berdiri, dan berjalan. Selain itu, mereka juga mempelajari keterampilan penggunaan panca indera dan komunikasi sosial (Rahmi & Hijriati, 2021). Bayi dalam fase *golden age* ini perlu dipantau dengan cermat oleh orang tua dan dukungan tenaga kesehatan untuk memastikan kesehatan bayi kedepannya.

Bayi yang baru lahir sangat rentan sehingga berisiko tinggi mengalami kematian, diperkirakan mencapai 50% dan 75% kematian bayi terjadi pada tahun pertama (UNICEF, 2017). Menurut *World Health Organization* (WHO),

kemungkinan atau risiko terjadinya kematian sebelum mencapai usia satu tahun bagi anak yang lahir pada tahun atau periode tertentu, merupakan definisi dari Angka Kematian Bayi (AKB), yang dinyatakan sebagai angka per 1000 kelahiran hidup (WHO, 2024).

Angka Kematian Bayi (AKB) merupakan salah satu indikator yang mencerminkan tingkat kesejahteraan masyarakat, yang dijadikan tolak ukur capaian program yang diimplementasikan oleh pemerintah untuk pembangunan kesehatan (Bugelli dkk., 2021). Berdasarkan Peraturan Menteri Kesehatan Republik Indonesia Nomor 21 Tahun 2020, penyebab utama kematian bayi yaitu disebabkan oleh gangguan pada masa perinatal (49,8%), kelainan kongenital dan genetik (14,2%), pneumonia (9,2%), diare dan infeksi gastrointestinal lainnya (7%), *viral hemorrhagic fever* (2,2%), meningitis (2%), serta gangguan undernutrisi dan metabolik (1,3%) (Kemenkes, 2020).

Kematian bayi masih menjadi masalah kesehatan yang serius di banyak negara, termasuk Indonesia. Data dari Badan Pusat Statistik (BPS) mengatakan bahwa selama 50 tahun terakhir, Angka Kematian Bayi (AKB) di Indonesia mengalami penurunan sekitar 90%. Penurunan tersebut tergolong signifikan. Berdasarkan Sensus Penduduk 2010, terdapat 26 kematian per 1000 kelahiran hidup, sedangkan hasil *Long Form* Sensus Penduduk 2020 menunjukkan penurunan menjadi 16,85 kematian per 1000 kelahiran hidup (BPS, 2023). Namun meskipun demikian, angka tersebut masih tergolong relatif tinggi dan belum mencapai Target 3.2 pada SDGs, yaitu kurang dari 12 per 1000 kelahiran hidup (SDGs Indonesia, 2023).

Berdasarkan hasil *Long Form* Sensus Penduduk 2020, pada tingkat regional, AKB tertinggi terdapat pada Provinsi Papua yaitu sebesar 38,17 per 1000 kelahiran hidup dan AKB terendah yaitu sebesar 10,38 per 1000 kelahiran hidup terdapat pada Provinsi DKI Jakarta. Angka tersebut menunjukkan bahwa terdapat kesenjangan AKB yang cukup besar antar provinsi di Indonesia. Secara statistik, AKB pada provinsi yang berada di dalam Pulau Jawa menunjukkan angka yang lebih rendah dibandingkan dengan AKB pada provinsi yang berada di luar Pulau Jawa (BPS, 2023).

Provinsi Jawa Timur merupakan salah satu provinsi di Pulau Jawa yang memiliki AKB tergolong rendah, yaitu sebesar 13,49 per 1000 kelahiran hidup. Selama 50 tahun terakhir, AKB pada Provinsi Jawa Timur juga mengalami penurunan yang cukup signifikan, namun angka tersebut tetap belum memenuhi target SDGs, yaitu masih di atas 12 per 1000 kelahiran hidup. Kesenjangan AKB masih banyak terjadi pada kabupaten/kota di Jawa Timur, dimana sebagian besar daerah tapal kuda (Banyuwangi, Bondowoso, Jember, Lumajang, Pasuruan, Situbondo dan Probolinggo) dan keempat kabupaten di Pulau Madura (Bangkalan, Sampang, Pamekasan, Sumenep) menunjukkan AKB yang relatif lebih tinggi dibandingkan kabupaten/kota lainnya (Diskominfo, 2023). Hal tersebut menandakan bahwa masih perlu dilakukan perbaikan untuk menekan tingginya Angka Kematian Bayi (AKB) di Provinsi Jawa Timur.

Berbagai peneliti telah melakukan penelitian mengenai kematian bayi, baik di tingkat nasional maupun internasional. Mulugeta dkk. (2022) telah melakukan pemodelan terhadap jumlah kematian bayi di Ethiopia menggunakan model multilevel log linear. Kusumawardani & Handayani (2018) meneliti tentang karakteristik ibu dan faktor risiko kejadian kematian bayi di Kabupaten Banjarnegara dengan penelitian survei berdasarkan pendekatan kasus kontrol dan dianalisis menggunakan uji *Chi-Square*. Sementara itu, Suhaeri & Sugiharti (2020) melakukan analisis mengenai pengaruh faktor sosial ekonomi terhadap Angka Kematian Bayi (AKB) di Provinsi Jawa Timur menggunakan regresi data panel.

Salah satu tantangan yang sering muncul dalam menganalisis data AKB adalah keberadaan pencilan. Keberadaan pencilan dapat disebabkan oleh kesalahan pengumpulan data, ketidakakuratan pada sistem pengukuran, variabilitas kondisi wilayah, bahkan fenomena tertentu seperti krisis ataupun bencana (Daniel, 2019). Pencilan didefinisikan sebagai amatan yang sangat berbeda atau menyimpang dari mayoritas data (Montgomery dkk., 2012). Keberadaan pencilan dalam model dapat berupa amatan pada variabel respon yang memiliki residual besar dan amatan pada variabel prediktor yang menyimpang dari pusat distribusi prediktor dalam model atau biasa disebut *leverage points* (Gibbert dkk., 2021). Keberadaan pencilan dalam model dapat menyebabkan penduga parameter menjadi bias dan ragam dari residual meningkat (Begashaw & Yohannes, 2020; Filzmoser & Nordhausen, 2021).

Selain pencilan, tantangan lain yang sering muncul saat menganalisis data AKB adalah multikolinearitas. Hal ini dikarenakan AKB dapat dipengaruhi oleh berbagai faktor sehingga memerlukan banyak variabel prediktor dalam memodelkannya. Banyaknya variabel prediktor tersebut akan menyebabkan data berdimensi tinggi dan berpotensi terjadi pelanggaran multikolinearitas (Nurfadilah dkk., 2024; Santi dkk., 2021). Multikolinearitas terjadi ketika terdapat korelasi linier yang tinggi antara variabel prediktor sehingga variabel prediktornya bersifat tidak saling bebas (Montgomery dkk., 2012). Hal ini dapat menyebabkan ragam pada model dan ragam penduga membesar sehingga penduga dapat menjadi tidak stabil (Khalaf & Iguernane, 2016).

Variabel Angka Kematian Bayi (AKB) termasuk variabel bertipe kontinu karena merepresentasikan tingkat atau *rate* dan diperoleh melalui proses pengukuran. Variabel kontinu adalah variabel yang dapat memiliki nilai dalam rentang tertentu tanpa batasan angka spesifik dan tidak terbatas pada kategori atau bilangan diskrit, termasuk bilangan bulat, desimal, atau pecahan (Guzik & Więckowska, 2023). Metode statistik yang umum digunakan untuk memodelkan variabel respon bertipe kontinu adalah regresi linier (Austin & Steyerberg, 2015).

Regresi linier merupakan salah satu algoritma *machine learning* yang paling sederhana dengan menggunakan pendekatan matematis untuk analisis prediktif. Regresi linier dapat digunakan untuk memodelkan hubungan antara variabel respon dengan variabel prediktornya (Maulud & Abdulazeez, 2020; Nurfadilah dkk., 2024). Regresi linier dapat berbentuk regresi linier sederhana ataupun regresi linier berganda, dimana perbedaan antar keduanya terletak pada banyaknya variabel prediktor. Regresi linier berganda dapat melibatkan lebih dari satu variabel prediktor dalam model (Maxwell dkk., 2019; Santi dkk., 2024).

Penduga parameter yang umum digunakan pada model regresi linier berganda adalah *Ordinary Least Squares* (OLS). Penduga OLS memiliki sifat-sifat penduga parameter yang baik, yaitu tak bias, efisien, dan konsisten. Hal ini dikarenakan OLS akan memiliki sifat BLUE (*Best Linear Unbiased Estimator*) ketika asumsi Gauss–Markov terpenuhi (Gujarati, 2003). Namun, penduga OLS pada model linier sangat sensitif terhadap pencilan. Adanya pencilan, bahkan hanya satu saja, dapat sangat memengaruhi hasil dugaan parameter (Begashaw & Yohannes, 2020; Shahzad dkk.,

2021). Hal tersebut dikarenakan OLS meminimumkan jumlah kuadrat galat dengan memberikan bobot yang sama untuk setiap amatan pada sampel. Namun, pencilan tidak akan memiliki dampak yang sama pada hasil regresi sehingga bobotnya tidak dapat diasumsikan sama. Pencilan dapat menarik garis regresi dan memiliki residual yang lebih besar sehingga dapat mendistorsi hasil analisis (Gujarati, 2003).

Metode alternatif yang dapat digunakan untuk mengatasi kekurangan OLS dalam menangani pencilan adalah regresi *robust* (Abdi dkk., 2020; Ayinde dkk., 2015). Regresi *robust* merupakan metode regresi alternatif yang efektif untuk menganalisis data yang dipengaruhi oleh pencilan karena menawarkan hasil yang lebih adaptif. Metode ini tetap memanfaatkan seluruh data, tetapi memberikan bobot lebih kecil pada data yang tergolong pencilan sehingga hasil pendugaan menjadi lebih resisten terhadap pencilan (Akolo & Nadjamuddin, 2022). Adapun jenis-jenis penduga *robust* yaitu *Least Absolute Deviation* (LAD), *Least Median Squares* (LMS), *Least Trimmed Squares* (LTS), *M-Estimation*, *S-Estimation*, dan *MM-Estimation* (Filzmoser & Nordhausen, 2021; Khan dkk., 2021).

Setiap penduga *robust* memiliki prinsip kerja yang berbeda-beda. Pendugaan parameter dari LAD diperoleh dengan meminimumkan jumlah nilai absolut dari galat. Namun penduga LAD kurang efektif dalam menangani pencilan dari variabel prediktor atau *leverage points* sehingga akan memiliki *breakdown point* dan efisiensi yang rendah. Sementara itu, penduga LMS meminimalkan median dari kuadrat galat dan memiliki *breakdown point* 50% sehingga lebih *robust* dibandingkan OLS yang memiliki *breakdown point* 0%. Namun, efisiensinya rendah yaitu 37%. Sama seperti LMS, penduga LTS juga memiliki *breakdown point* 50%, namun efisiensinya sangat rendah karena prinsipnya meminimumkan jumlah kuadrat galat yang telah dipangkas atau *trimmed* (Begashaw & Yohannes, 2020; Filzmoser & Nordhausen, 2021; Khan dkk., 2021).

M-Estimation mengatasi kekurangan penduga LAD, LMS, dan LTS terkait efisiensi rendah. Berbeda dengan OLS yang menggunakan jumlah kuadrat galat sebagai fungsi *loss*, *M-Estimation* menggunakan fungsi *loss* tertentu yang dirancang untuk membatasi pengaruh residual besar. Pendekatan ini memungkinkan *M-Estimation* menjadi lebih *robust* terhadap pencilan dan tetap mempertahankan efisiensi yang tinggi, yaitu sekitar 95%. Namun, *M-Estimation*

memiliki kelemahan seperti LAD, yaitu kurang efektif dalam menangani pencilan dari variabel prediktor (*leverage points*) sehingga memiliki *breakdown point* yang rendah. Sementara itu, *S-Estimation* dapat menghasilkan *high breakdown point* 50%, namun dengan efisiensi yang rendah (Khan dkk., 2021). Yohai (1987) mengembangkan *MM-Estimation* dengan menggabungkan *S-Estimation* dan *M-Estimation* sehingga memiliki *high breakdown point* 50% dan efisiensi tinggi mendekati 95%. *MM-Estimation* merupakan satu-satunya penduga *robust* yang memiliki dua sifat ini secara bersamaan (Ayinde dkk., 2015).

Oyeyemi dkk. (2021) telah membandingkan beberapa penduga *robust* seperti *M-Estimation*, *MM-Estimation*, *S-Estimation*, *LTS*, *LMS*, dan *LAD* dalam regresi linier berganda yang mengandung pencilan pada data simulasi. Hasil penelitian menunjukkan bahwa *MM-Estimation* memiliki ragam yang lebih kecil dan tetap stabil saat ukuran sampel bertambah, terlepas dari tingkat ragam pencilannya. Selain itu, Sari dkk. (2020) melakukan perbandingan antara regresi linier berganda dengan penduga *Ordinary Least Squares* (*OLS*) dan *MM-Estimation* untuk kasus Demam Berdarah Dengue (*DBD*) di Indonesia. *MM-Estimation* terbukti lebih efisien dalam menganalisis model yang mengandung pencilan dibandingkan *OLS*.

MM-Estimation dapat mengatasi kekurangan *OLS* dalam menangani masalah pencilan, namun metode ini masih belum mampu menangani pelanggaran multikolinearitas. Terdapat berbagai metode yang dapat digunakan untuk mengatasi pelanggaran multikolinieritas, salah satunya yaitu *penalized regression* (Nurfadilah dkk., 2024). *Penalized regression* dapat mengatasi kekurangan *OLS* ketika terdapat pelanggaran multikolinearitas dengan cara menambahkan penalti pada fungsi objektif *OLS* (Abram dkk., 2016). *Penalized regression* memiliki beberapa pendekatan, diantaranya yaitu *Least Absolute Shrinkage and Selection Operator* (*LASSO*), *Minimax Concave Penalty* (*MCP*), *Smoothly Clipped Absolute Deviation* (*SCAD*), *Elastic net*, dan *Ridge Regression* (Santi dkk., 2019; Santi dkk., 2021; Wang dkk., 2022).

Ridge Regression atau regresi *ridge* merupakan metode regresi alternatif yang dikembangkan oleh Hoerl & Kennard (1970) sebagai perluasan dari metode *OLS*. Dalam pendugaan parameter, regresi *ridge* meminimumkan jumlah kuadrat galat dengan menambahkan penalti L_2 berbentuk kuadrat sehingga menghasilkan

penduga yang bersifat bias, namun lebih stabil (Nurfadilah dkk., 2024). Regresi *ridge* itu sendiri terbagi menjadi dua, yaitu *Ordinary Ridge Regression* (ORR) dan *Generalized Ridge Regression* (GRR). Dalam regresi *ridge*, terdapat parameter tambahan yang berperan penting yaitu konstanta bias atau penalti *ridge*. Perbedaan antara ORR dan GRR terletak pada besaran penalti tersebut, dimana konstanta bias pada GRR dapat berbeda untuk setiap variabel prediktor sehingga lebih fleksibel karena tidak memaksakan besaran penalti yang sama (Fayose & Ayinde, 2019).

Permasalahan menjadi lebih kompleks ketika masalah pencilan dan multikolinearitas terjadi secara bersamaan. Regresi *robust* hanya dapat menangani masalah pencilan, sedangkan regresi *ridge* hanya dapat menangani pelanggaran multikolinearitas. Kedua metode tersebut tidak dapat menangani masalah pencilan dan pelanggaran multikolinearitas secara bersamaan (Adegoke dkk., 2016; Affindi dkk., 2019). Berbagai peneliti telah mengembangkan metode regresi *ridge robust*, yang mengombinasikan regresi *ridge* dengan regresi *robust*, salah satunya yaitu Samkar & Alpu (2010). Metode ini dapat mengatasi masalah pencilan dan multikolinearitas secara bersamaan. Regresi *ridge robust* dapat diperoleh melalui kombinasi berbagai penduga *robust* dengan ORR ataupun GRR.

Penelitian mengenai regresi *ridge robust* telah dilakukan oleh beberapa peneliti, seperti Jeremia dkk. (2020) yang membandingkan *Ordinary Ridge Regression* (ORR) serta *Generalized Ridge Regression* (GRR) dengan penduga *robust* yang sama, yaitu *S-Estimation*. Hasil penelitian menunjukkan bahwa kombinasi GRR dengan *S-Estimation* merupakan model terbaik pada penelitian tersebut dibandingkan kombinasi ORR dengan *S-Estimation*. Sementara itu, Destiyani dkk. (2019) menerapkan regresi *ridge robust* dengan kombinasi ORR dan *MM-Estimation* pada Angka Kematian Bayi (AKB) di Provinsi Jawa Tengah.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka rumusan masalah untuk penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan *Generalized Ridge Regression* dengan *MM-Estimation* (GRR – MM) pada Angka Kematian Bayi (AKB) di Provinsi Jawa Timur?
2. Faktor apa saja yang memiliki pengaruh secara signifikan terhadap AKB di Provinsi Jawa Timur dari hasil penerapan GRR – MM?

1.3. Batasan Masalah

Penelitian ini masih terdapat beberapa batasan masalah diantaranya adalah sebagai berikut:

1. Pencilan yang dianalisis adalah pencilan dalam model, bukan secara univariat.
2. Fungsi pembobot yang digunakan adalah fungsi pembobot *Tukey Bisquare*.
3. Penentuan konstanta bias hanya menggunakan metode Hoerl dan Kennard.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah dan batasan masalah yang telah dipaparkan sebelumnya, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan *Generalized Ridge Regression* (GRR) dengan *MM-Estimation* (GRR – MM) pada Angka Kematian Bayi (AKB) di Provinsi Jawa Timur.
2. Menganalisis faktor yang berpengaruh secara signifikan terhadap AKB di Provinsi Jawa Timur berdasarkan penerapan GRR – MM.

1.5. Manfaat Penelitian

Hasil penelitian ini dapat menambah wawasan khususnya pada bidang ilmu statistika mengenai pemodelan regresi untuk mengatasi masalah pencilan dan pelanggaran multikolinearitas. Selain itu, penelitian ini juga dapat menjadi acuan untuk memberikan kontribusi bagi pemerintah serta pihak terkait di Provinsi Jawa Timur dalam menentukan strategi yang efektif dan efisien untuk mencapai Target 3.2 pada SDGs.