

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan ilmu statistika dan analisis semakin pesat pada era modern saat ini. Salah satu metode statistika yang umum digunakan adalah regresi linear sederhana yang mampu menjelaskan hubungan antara satu variabel bebas dengan satu variabel terikat serta mengestimasi pengaruh variabel bebas terhadap variabel terikat. Penerapan metode ini digunakan secara luas dalam berbagai bidang ilmu untuk mendukung proses pengambilan keputusan yang didasarkan pada analisis data. Menurut Montgomery et al. (2021), regresi linear sederhana merupakan salah satu metode utama dalam statistika karena memberikan gambaran teori yang kuat untuk memodelkan hubungan antar variabel, serta kerangka *linear models* telah digunakan secara luas di berbagai bidang ilmu pengetahuan (Castro Torres & Akbaritabar, 2024). Oleh karena itu, pemahaman terhadap konsep dasar dan asumsi-asumsi yang mendasari model regresi linear sederhana menjadi aspek penting bagi peneliti dalam penggunaan metode analisis.

Secara umum, model regresi linear sederhana dinyatakan dalam bentuk

$$Y = X\beta + \varepsilon$$

dimana Y merupakan variabel terikat dengan ukuran $(n \times 1)$. Matriks X terdiri atas konstanta dan satu variabel bebas sehingga berukuran $(n \times 2)$. Vektor parameter $\beta = (\beta_0, \beta_1)^T$ berukuran (2×1) , dimana β_0 dan β_1 masing-masing merupakan intersep dan kemiringan garis (*slope*). Sementara itu, ε adalah galat (*error*) yang bersifat acak berukuran $(n \times 1)$. Model regresi linear sederhana didasarkan pada beberapa asumsi klasik terhadap galat, yaitu bersifat independensi galat, homoskedastisitas, dan normalitas distribusi galat (Draper & Smith, 1998), dan pemenuhan asumsi-asumsi tersebut juga diperkuat dalam penelitian terkini (Yang et al., 2019). Asumsi normalitas distribusi *error* secara khusus dianggap memegang peran penting karena berkaitan langsung dengan teknik uji statistik, seperti pengujian parameter model dan pembentukan interval kepercayaan (Shatz, 2024; Wooldridge, 2020). Gujarati & Porter, (2009) menyatakan bahwa apabila terjadi pelanggaran asumsi normalitas dapat mengurangi validitas uji statistik, khususnya pada ukuran sampel kecil. Hal ini juga didukung oleh penelitian terbaru

yang menunjukkan bahwa pelanggaran asumsi tersebut berdampak nyata terhadap hasil analisis dalam regresi linear klasik (Rajh-Weber et al., 2025).

Namun, dalam praktiknya, data empiris tidak selalu memenuhi asumsi bahwa *error* berdistribusi normal. Pada bidang pendidikan misalnya, asumsi normalitas *error* sering kali kurang realistis untuk menggambarkan kondisi data yang sebenarnya (Blanca et al., 2017; Wooldridge, 2020). Secara teoritis, distribusi normal memiliki rentang nilai yang tidak terbatas, yaitu $(-\infty, \infty)$, sehingga memungkinkan munculnya nilai galat yang sangat besar maupun sangat kecil (Hogg et al., 2019). Sementara itu, variabel terikat dalam konteks tertentu seperti hasil belajar umumnya memiliki batas nilai tertentu, misalnya dalam rentang 0 sampai 100 (Arthurs et al., 2019; Fraenkel et al., 2012; Nitko & Brookhart, 2019). Dengan mempertimbangkan karakteristik tersebut, asumsi mengenai distribusi galat perlu disesuaikan agar sesuai dengan batas pengukuran variabel terikat. Salah satu pendekatan yang dapat digunakan adalah dengan menganggap galat mengikuti distribusi *uniform* pada interval $(-a, a)$, sehingga nilai galat berada dalam batas tertentu. Pembatasan ini memungkinkan adanya variasi acak dalam model tanpa mengabaikan bahwa variabel terikat memiliki rentang nilai yang terbatas.

Sejumlah penelitian sebelumnya telah mengkaji analisis regresi linear dengan asumsi *error non-normal*. Salah satunya penelitian terdahulu dilakukan oleh Jlibene et al. yang membahas karakteristik dari estimator kuadrat terkecil (*Least Square Estimator* / LSE) dalam model regresi linear sederhana dengan asumsi *error* berdistribusi *uniform* $(-\theta, \theta)$ (Jlibene et al., 2021). Hasil penelitian membuktikan bahwa meskipun asumsi normalitas tidak terpenuhi, estimator LSE tetap bersifat *Best Linear Unbiased Estimator* (BLUE) dan teorema Gauss–Markov tetap berlaku (Jlibene et al., 2021). Penelitian lain oleh Wang yang membandingkan performa *Least Squares Estimator* (LSE) dan *Maximum Likelihood Estimator* (MLE) dalam analisis regresi linear sederhana dengan *error* berdistribusi *non-normal* yaitu pada distribusi *uniform* dan *laplace* (Wang, 2025). Hasil penelitian dengan pendekatan simulasi *monte carlo* menunjukkan bahwa MLE cenderung memiliki *Mean Squared Error* (MSE) dan variansi yang lebih kecil dibandingkan LSE, terutama ketika ukuran sampel meningkat. Namun pada ukuran sampel kecil, LSE dapat memiliki bias yang lebih kecil dibandingkan MLE (Wang, 2025).

Kedua penelitian tersebut memberikan relevansi yang signifikan dalam pengembangan analisis regresi dengan *error* berdistribusi *non-normal*. Namun, keduanya masih memiliki keterbatasan yang dapat dikembangkan dalam penelitian lebih lanjut. Jlibene et al. (2021) telah mengkaji estimator kuadrat terkecil pada regresi linear dengan galat berdistribusi *uniform*, namun belum secara mendalam mengkaji kesesuaian hasil estimasi yang diperoleh dari metode kuadrat terkecil dan metode maksimum *likelihood* berdasarkan pendekatan simulasi. Sementara itu, Wang (2025) menitikberatkan pada pendekatan simulasi tanpa mengkaji secara mendalam sifat-sifat teoritis estimator dalam distribusi *uniform*. Oleh karena itu, penelitian ini dilakukan untuk melengkapi kajian terdahulu dengan menganalisis secara teoritis bentuk estimasi parameter, mengkaji karakteristik estimator parameter, serta mengevaluasi kesesuaian antara hasil analisis teoritis dan hasil simulasi dalam estimasi parameter regresi linear sederhana dengan galat berdistribusi *uniform* pada interval $(-a, a)$.

1.2 Rumusan Masalah

Berdasarkan pembahasan pada latar belakang, dapat dirumuskan permasalahan yang menjadi dasar dalam penelitian ini sebagai berikut:

1. Bagaimana merumuskan bentuk estimasi parameter regresi (β_0, β_1) serta parameter distribusi galat a pada model regresi linear sederhana dengan galat berdistribusi *uniform*?
2. Bagaimana menganalisis sifat-sifat estimator parameter yang diperoleh?
3. Bagaimana kesesuaian antara hasil teoritis dan hasil simulasi dalam estimasi parameter pada model regresi linear sederhana dengan galat berdistribusi *uniform*?

1.3 Batasan Masalah

Berdasarkan rumusan masalah yang telah dirumuskan, maka untuk memperjelas fokus penelitian ini, ditetapkan batasan masalah sebagai berikut:

1. Asumsi galat dibatasi pada distribusi *uniform*, tanpa mempertimbangkan distribusi lain.
2. Analisis difokuskan pada sifat-sifat estimator parameter, tanpa membahas interval estimasi dan uji hipotesis.

1.4 Tujuan Penelitian

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

1. Mendapatkan perumusan bentuk estimasi parameter regresi (β_0, β_1) serta parameter distribusi galat a .
2. Menganalisis sifat-sifat estimator parameter yang diperoleh.
3. Menganalisis kesesuaian antara hasil teoritis dan hasil simulasi dalam estimasi parameter pada model regresi linear sederhana dengan galat berdistribusi *uniform*.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memberikan pemahaman bahwa regresi linear sederhana tetap dapat dianalisis meskipun galat tidak berdistribusi normal.
2. Menjadi referensi dalam pengembangan penelitian statistika terkait estimator parameter pada kondisi galat *non-normal*.
3. Memberikan wawasan praktis mengenai pentingnya kesesuaian asumsi distribusi galat dalam pemodelan regresi.

