

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis regresi seringkali digunakan untuk mengetahui bentuk hubungan antara variabel respon dengan satu atau lebih variabel prediktor. Analisis regresi yang umumnya digunakan adalah analisis regresi klasik, dimana variabel responnya merupakan data kontinu yang mengikuti distribusi normal. Namun dalam perkembangannya model regresi klasik ini tidak mampu mengatasi permasalahan-permasalahan dimana variabel respon berupa data diskrit dan tidak berdistribusi normal.

Solusi untuk mengatasi masalah tersebut adalah dengan menggunakan model linier tergeneralisir atau biasa disebut *Generalized Linear Model (GLM)*. Uji asumsi yang diterapkan pada GLM tidak mengharuskan asumsi kenormalan dari variabel respon dan juga tidak mengharuskan kehomogenan dari variansinya (de Jong dan Heller, 2008). Salah satu model regresi yang termasuk dalam penerapan GLM adalah regresi Poisson dimana dalam regresi ini variabel respon diasumsikan berdistribusi Poisson, dengan fungsi massa peluangnya adalah

$$f(y; \mu) = \frac{e^{-\mu} \mu^y}{y!}$$

Regresi Poisson digunakan untuk menganalisis hubungan antara variabel respon yang berupa data diskrit dengan satu atau lebih variabel independen, dimana nilai mean dan variansinya diasumsikan sama (equidispersi).

$$E(Y) = Var(Y) = \mu$$

Pada prakteknya seringkali data cacah memperlihatkan variasi yang sangat besar dimana variansi sampel lebih besar daripada mean sampel (overdispersi) atau sebaliknya mean sampel lebih besar daripada variansi sampel (underdispersi). Apabila regresi Poisson diaplikasikan untuk data dispersi (overdispersi/ underdispersi) maka hasil analisisnya tidak tepat karena hubungan antara variansi dan mean pada regresi Poisson yang mengalami dispersi adalah

$$Var(Y) = \phi E(Y) = \phi \mu$$

dimana  $\phi$  adalah parameter dispersi/ *scale parameter*. Ketika terjadi overdispersi pada regresi Poisson maka  $\phi$  bernilai lebih dari satu dan bersifat konstan.

Ada beberapa metode yang dapat digunakan untuk mengatasi overdispersi pada regresi Poisson. Dalam Tugas Akhirnya Rio T. Simarmata (2010) membahas mengenai regresi binomial negatif, model regresi ini dapat mengatasi overdispersi pada regresi Poisson karena variabel responnya diasumsikan mengikuti distribusi binomial negatif sehingga tidak mengharuskan nilai mean yang sama dengan nilai variansi. Selain regresi binomial negatif metode lain untuk mengatasi overdispersi regresi Poisson adalah model regresi zero-inflated Poisson, metode ini digunakan jika overdispersi disebabkan oleh banyaknya variabel respon yang bernilai nol. Pada Tugas Akhir ini kasus overdispersi regresi Poisson akan diatasi dengan menggunakan regresi quasi-likelihood. Regresi quasi-likelihood dapat digunakan untuk kasus overdispersi karena model ini mengasumsikan  $Var(Y_i) = \phi \mu_i$ . Metode ini terdiri dari tiga tahapan yakni mengestimasi parameter regresi, mengestimasi parameter dispersi  $\phi$ , dan menyesuaikan standar error dengan nilai estimasi parameter dispersi ( $\hat{\phi}$ ).

## 1.2 Perumusan Masalah

Permasalahan yang akan dibahas dalam Tugas Akhir ini adalah bentuk model regresi quasi-likelihood, estimasi parameter, analisis kecocokan model, dan signifikansi koefisien regresi, serta aplikasinya dalam mengatasi overdispersi regresi Poisson.

## 1.3 Pembatasan Masalah

Permasalahan dibatasi pada estimasi parameter regresi quasi-likelihood dengan menggunakan *Quasi-Likelihood Estimation* (QLE) melalui *Iterative Reweighted Least Square* (IRWLS) dan estimasi parameter dispersi menggunakan nilai deviansi dan Pearson chi-square.

## 1.4 Tujuan Penulisan

1. Mengetahui terjadinya overdispersi pada regresi Poisson
2. Menentukan model regresi quasi-likelihood
3. Mengestimasi parameter model regresi quasi-likelihood dan parameter dispersi.
4. Menentukan statistik uji kecocokan model dan statistik uji koefisien regresi.

## 1.5 Manfaat Penulisan

Manfaat penulisan Tugas Akhir ini adalah memperoleh alternatif solusi yang dapat digunakan untuk mengatasi overdispersi pada regresi Poisson.

## **1.6 Sistematika Penulisan**

Sistematika penulisan pada tugas akhir ini meliputi empat bab. BAB I merupakan latar belakang, permasalahan, pembatasan masalah, tujuan penulisan dan sistematika penulisan. BAB II merupakan teori penunjang yang akan digunakan sebagai acuan dalam memahami permasalahan yang akan dibahas. BAB III merupakan pembahasan yang berisi bentuk model, penaksiran parameter, pengujian signifikansi model dan parameter serta contoh aplikasi dari model regresi quasi-likelihood. BAB IV merupakan hasil yang dapat diambil dari penulisan tugas akhir.