

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah suatu metode yang digunakan untuk menganalisis hubungan antara variabel respon dengan beberapa variabel prediktor. Pada umumnya analisis regresi digunakan untuk menganalisis data variabel respon yang berupa data kontinu. Namun dalam beberapa aplikasinya, data variabel respon yang akan dianalisis dapat berupa data diskrit. Salah satu model regresi yang dapat digunakan untuk menganalisis hubungan antara variabel respon Y yang berupa data diskrit dengan variabel prediktor X data diskrit, kontinu, kategorik atau campuran adalah model regresi poisson.

Dalam model regresi poisson terdapat beberapa asumsi yang harus dipenuhi, salah satunya adalah nilai variansi dari variabel respon Y yang diberikan oleh $X = x$ harus sama dengan nilai meannya yaitu:

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

Namun dalam analisis data diskrit dengan menggunakan model regresi poisson terkadang terjadi pelanggaran asumsi tersebut, dimana nilai variansinya lebih besar dari nilai mean yang disebut overdispersi atau varian lebih kecil dari nilai mean yang disebut underdispersi. Dalam model regresi linier klasik pelanggaran tersebut dinamakan pelanggaran asumsi homokedastisitas (Cameron dan Trivedi, 1998). Jika terjadi kasus overdispersi, penanganan model yang dapat digunakan adalah model regresi binomial negatif. Model ini dibentuk dari distribusi *mixture* poisson-gamma dimana :

$$y | \mu \sim \text{poisson}(\mu)$$

$$\mu \sim \text{Gamma}(\alpha, \beta)$$

Jika suatu distribusi poisson (μ) dimana μ merupakan nilai variabel random yang berdistribusi gamma maka akan dihasilkan distribusi mixture yang dinamakan distribusi binomial negatif.

Model regresi binomial negatif mengasumsikan

$$E(Y | x) = \mu \quad \text{dan} \quad \text{Var}(Y | x) = \mu + k\mu^2$$

k adalah parameter dispersi = $1/\alpha$ dimana α = parameter bentuk dari distribusi gamma. Model ini dapat mengatasi masalah overdispersi karena tidak mengharuskan nilai mean yang sama dengan nilai variansi seperti pada model regresi poisson.

1.2 Permasalahan

Dalam penulisan tugas akhir ini, permasalahan yang akan dibahas adalah bentuk model regresi binomial negatif, estimasi parameter, analisis kecocokan model dan signifikansi koefisien regresi binomial negatif.

1.3 Pembatasan Masalah

Penulisan tugas akhir ini dibatasi pada estimasi parameter regresi binomial negatif dan parameter dispersi dengan menggunakan metode maksimum likelihood melalui prosedur iterasi Newton-Rhapson dan Fisher Scoring.

1.4 Tujuan penulisan

Tujuan penulisan tugas akhir ini adalah :

1. Mendiagnosa overdispersi pada model regresi poisson
2. Menentukan model regresi binomial negatif
3. Mengestimasi parameter model regresi binomial negatif dan parameter dispersi.
4. Menentukan statistik uji kecocokan model dan statistik uji koefisien regresi.

1.5 Manfaat Penulisan

Manfaat penulisan tugas akhir ini adalah dapat digunakan untuk menentukan model regresi yang digunakan untuk menganalisis hubungan antara variabel respon bertipe diskrit dengan beberapa variabel prediktor, untuk selanjutnya dapat diuji kontribusi ataupun signifikansi variabel prediktor terhadap model regresi tersebut.

1.6 Sitematika 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 terdiri dari : fungsi distribusi peluang, distribusi binomial, distribusi poisson, distribusi gamma, distribusi binomial negatif, model regresi linier, generalisasi model linier, metode maksimum likelihood, model regresi poisson, overdispersi regresi poisson. BAB III merupakan pembahasan yang berisi bentuk model, penaksiran parameter, pengujian signifikansi model dan parameter serta contoh aplikasi dari model regresi binomial negatif. BAB IV merupakan hasil: yang dapat diambil dari penulisan tugas akhir.