

INTISARI

MODEL REGREI *FINITE MIXTURE* BINOMIAL NEGATIF

Oleh

Taufiq Dwi Apriyanto
12/334605/PA/14838

Model campuran merupakan model gabungan dari dua atau lebih distribusi. Distribusi yang digunakan dalam model campuran bisa distribusi yang sama dengan parameter berbeda maupun distribusi yang berbeda. Model ini mengelompokkan data di dalam suatu dataset menjadi kelompok-kelompok data yang sebelumnya tidak terdefinisikan. Salah satu model campuran yang dikembangkan adalah model regresi *finite mixture* Binomial Negatif. Model ini dapat mengatasi adanya overdispersi pada regresi Poisson sekaligus mengatasi adanya heterogenan pada data.

Estimasi parameter model regresi *finite mixture* Binomial Negatif diestimasi dengan metode *Maximum Likelihood Estimator* (MLE) yang menggunakan algoritma Ekspektasi Maksimalisasi (EM) untuk mendapatkan parameter MLE-nya. Pemilihan model terbaik dilakukan dengan cara membandingkan nilai BIC dan AIC dengan model regresi cacah lainnya.

Model regresi *finite mixture* Binomial Negatif diaplikasikan pada data tentang jumlah ketidakhadiran siswa sekolah menengah pada mata pelajaran Bahasa Portugis di Negara Portugal. Hasilnya menunjukkan model ini memiliki AIC dan BIC terkecil.

Kata kunci: overdispersi, heterogen, model campuran, *finite mixture*, binomial negatif.

ABSTRACT

FINITE MIXTURE NEGATIVE BINOMIAL REGRESSION MODEL

by

Taufiq Dwi Apriyanto
12/334605/PA/14838

Mixture model is a combination model of two or more distributions. Distribution in the model mixture can consist of distribution identic with different parameter or non-identic. This model classify data within dataset into groups of data that previously were not defined. One of the models is finite mixture Negative Binomial regression model. This model can handle overdispersion on Poisson regression and heterogeneity in the data.

Parameter estimation for finite mixture Negative Binomial regression model is estimated by Maximum Likelihood Estimator (MLE) which uses Expectation Maximization (EM) algorithm for getting the MLE's parameter. The best model selection is evaluated by comparing the AIC and BIC score with other discrete regression models.

Finite mixture Negative Binomial regression model is applied on number of absences from middle school students on the subjects of Portuguese language in the State Portugal. The result showed this model has the lowest AIC and BIC.

Keywords: overdispersion, heterogeneity, mixture model, finite mixture, negative binomial