

ABSTRACT

Bayesian Model Selection in Quantile Regression with Bayes Factor

by

Angelia Trince Monekaka
14/366174/PA/16220

Quantile regression is used to overcome the limitation of linear regression in analyzing the unfulfilled assumptions of classical regression. The interest focuses on relationship between predictors and the quantile of the distribution of a continuous response.

Quantile regression can be estimated using Bayesian method which is based on information from sample and prior information. Combination of those information is called posterior. Gibbs sampling approach is used to find the posterior distribution. In Bayesian quantile regression, there is uncertainty in which of the many candidate predictors should be included. In order to identify important predictors and to build accurate predictive model, Bayesian method for variable selection and model selection are very useful.

The case study in this thesis aims to predict body weight according to the 12 variables of body girth measurements and height. Models from Bayesian variable selection are used as candidate models Bayes factor for Bayesian model selection. Model comparison with Bayes factor shows quantify the relative evidence for one over another model. Therefore, model selection with Bayes factor is model with 9 parameters (C, ChestGi, WaistGi, HipGi, ThighGi, ForearmGi, KneeGi, CalfGi, Height) shows that data are about 43,741 times more likely under this model than under the other alternative models.

Keywords: *Quantile Regression, Bayesian, Gibbs sampling, Bayesian model selection, Bayes Factor*

INTISARI

Seleksi Model Regresi Kuantil Bayesian dengan Bayes *Factor*

Oleh

Angelia Trince Monekaka
14/366174/PA/16220

Regresi kuantil digunakan untuk mengatasi keterbatasan regresi linear dalam menganalisis asumsi yang tidak terpenuhi pada regresi klasik. Hal ini berfokus pada hubungan antara variabel prediktor dan kuantil dari distribusi variabel dependen yang kontinu.

Regresi kuantil diestimasi dengan metode Bayesian yang berdasarkan pada informasi yang berasal dari sampel dan informasi prior. Gabungan informasi ini disebut posterior. Untuk mencari distribusi posterior digunakan pendekatan Gibbs *sampling*. Dalam analisis regresi kuantil Bayesian, terdapat ketidakpastian dari banyak kandidat prediktor yang harus dimasukkan dalam model. Untuk mengidentifikasi prediktor penting dan membangun model prediksi yang akurat, metode Bayesian untuk pemilihan variabel dan seleksi model sangat berguna.

Studi kasus yang digunakan dalam skripsi ini bertujuan untuk memprediksi berat badan (*Weight*) yang sesuai dengan 12 variabel pengukuran lingkar tubuh (*body girth measurements*) dan tinggi badan (*Height*). Model dari hasil seleksi variabel Bayesian digunakan sebagai kandidat model Bayes *factor* untuk seleksi model Bayesian. Perbandingan model dengan Bayes *factor* menunjukkan pengukuran bukti relatif untuk satu model di atas model lainnya. Dengan perbandingan model Bayes *factor*, didapatkan kesimpulan bahwa model dengan 9 parameter (C, ChestGi, WaistGi, HipGi, ThighGi, ForeaGi, KneeGi, CalfGi, *Height*) dengan data menjelaskan 43,741 kali lebih baik dibandingkan model alternatif lain.

Kata kunci: Regresi Kuantil, Bayesian, Gibbs *sampling*, Seleksi Model Bayesian, Bayes Factor