

Penaksiran parameter pada regresi kuantil bayesian = Parameter estimation of bayesian quantile regression

Dickson Dichandra, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20510511&lokasi=lokal>

Abstrak

Regresi kuantil adalah metode regresi yang menghubungkan kuantil dari variabel respon dengan satu atau beberapa variabel prediktor. Regresi kuantil memiliki kelebihan yang tidak dimiliki oleh regresi linier yaitu robust terhadap outlier dan dapat memodelkan data yang heteroskedastisitas. Regresi kuantil dapat diestimasi parameternya dengan metode Bayesian.

Metode Bayesian adalah alat analisis data yang diturunkan berdasarkan prinsip inferensi Bayesian. Inferensi Bayesian adalah proses mempelajari analisis data secara induktif dengan teorema Bayes. Untuk menaksir parameter regresi dengan inferensi Bayesian, perlu dicari distribusi posterior dari parameter regresi dimana distribusi posterior proporsional terhadap perkalian distribusi prior dan fungsi likelihoodnya. Karena perhitungan distribusi posterior secara analitik sulit untuk dilakukan jika semakin banyak parameter yang ditaksir, maka diajukan metode Markov Chain Monte Carlo (MCMC). Penggunaan metode Bayesian dalam regresi kuantil memiliki kelebihan yaitu penggunaan MCMC memiliki kelebihan yaitu mendapatkan sampel nilai parameter dari distribusi posterior yang tidak diketahui, penggunaan

yang efisien secara komputasi, dan mudah diimplementasikannya. Yu dan Moyeed (2001) memperkenalkan regresi kuantil Bayesian dengan menggunakan fungsi likelihood dari error yang berdistribusi Asymmetric Laplace Distribution (ALD) dan menemukan bahwa

meminimumkan taksiran parameter pada regresi kuantil sama dengan memaksimalkan fungsi likelihood dari error yang berdistribusi Asymmetric Laplace Distribution (ALD). Metode yang digunakan untuk menaksir parameter regresi kuantil adalah Gibbs sampling dari distribusi ALD yang merupakan kombinasi dari distribusi eksponensial dan Normal. Penaksiran parameter model regresi dilakukan dengan cara pengambilan sampel pada distribusi posterior

dari parameter regresi yang ditemukan dalam skripsi ini. Pengambilan sampel pada distribusi posterior dapat menggunakan metode Gibbs sampling. Hasil yang diperoleh dari Gibbs sampling berupa barisan sampel parameter yang diestimasi. Setelah mendapatkan barisan sampel, barisan sampel dirata-ratakan untuk mendapatkan taksiran parameter regresinya. Studi kasus dalam skripsi ini adalah membahas pengaruh faktor risiko dari nasabah asuransi kendaraan bermotor terhadap besar klaim yang diajukan oleh nasabah.

.....Quantile regression is a regression method that links the quantiles of the response variable with one or more predictor variables. Quantile regression has advantages that linear regression does not have; it is robust against outliers and can model heteroscedasticity data.

The parameters of quantile regression can be estimated using the Bayesian method. The Bayesian method is a data analysis tool derived based on the Bayesian inference principle.

Bayesian inference is the process of studying data analysis inductively with the Bayes theorem. To estimate regression parameters with Bayesian inference, it is necessary to find the posterior distribution of the regression parameters where the posterior distribution is

proportional to the product of the prior distribution and its likelihood function. Since the calculation of the

posterior distribution analytically is difficult to do if the more parameters are estimated, the Markov Chain Monte Carlo (MCMC) method is proposed. The use of the Bayesian method in quantile regression has advantages, namely the use of MCMC has the advantages of obtaining sample parameter values from an unknown posterior distribution, using computationally efficient, and easy to implement. Yu and Moyeed (2001) introduced Bayesian quantile regression using the likelihood function of errors with an Asymmetric Laplace Distribution (ALD) distribution and found that minimizing parameter estimates in quantile regression is the same as maximizing the likelihood function of errors with an Asymmetric Laplace Distribution (ALD) distribution. The method used to estimate quantile regression parameters is Gibbs sampling from the ALD distribution, which is a combination of the exponential and normal distributions. The estimation of the regression model parameters is done by sampling the posterior distribution of the regression parameters which is found in this thesis. Gibbs sampling method is used to sampling the posterior distribution. The results obtained from Gibbs sampling are a sample sequence of estimated parameters. After obtaining the sample sequences, the sample lines are averaged to obtain an estimated regression parameter. The case study in this thesis discusses the effect of risk factors from motor vehicle insurance customers on the size of claims submitted by customers.