

# Konstruksi model Zero-Inflated Ordered Probit untuk prediksi tingkat keparahan cedera akibat kecelakaan lalu lintas = Construction of Zero-Inflated Ordered Probit model for predicting injury severity due to traffic accident

Putri Maharani Ardiasih, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920547234&lokasi=lokal>

---

## Abstrak

Analisis regresi merupakan metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan satu atau lebih variabel prediktor. Model regresi yang sering digunakan adalah model regresi linier dengan asumsi variabel respon berdistribusi normal. Pengembangan dari model regresi linier adalah Generalized Linear Model (GLM). Salah satu komponen dari GLM adalah fungsi penghubung yang digunakan untuk menghubungkan variabel respon dengan prediktor linier. Pemilihan fungsi penghubung ini bergantung pada jenis variabel respon. Pada variabel respon kategorik di mana tidak ada keterurutan (nominal), salah satu GLM yang dapat digunakan adalah model dengan fungsi penghubung probit. Namun, model ini tidak dapat digunakan untuk variabel respon kategorik dengan keterurutan (ordinal). Untuk itu, dikembangkan model ordered probit untuk menganalisis variabel respon ordinal dengan fungsi penghubung probit. Namun, jika kategori 0 memiliki observasi yang lebih banyak dibandingkan kategori lainnya, maka terjadi zero-inflation pada data respon. Hal ini dapat menyebabkan terjadinya overdispersi yang berakibat pada kesalahan interpretasi model dan kesalahan pengambilan kesimpulan. Untuk itu, diperlukan pengembangan dari model ordered probit, yaitu model zero-inflated ordered probit. Estimasi parameter model zero-inflated ordered probit dilakukan menggunakan metode maximum likelihood. Implementasi dari model zero-inflated ordered probit digunakan untuk memprediksi tingkat keparahan cedera akibat kecelakaan lalu lintas berdasarkan data kecelakaan lalu lintas. Sebagai variabel respon ordinal pada model zero-inflated ordered probit adalah tingkat keparahan cedera yang memiliki tiga kategori, yaitu tidak terjadi cedera berat, cedera serius, dan cedera fatal. Pada data tersebut, kategori tidak terjadi cedera berat (0) memiliki observasi yang lebih banyak dibandingkan kategori lainnya sehingga terjadi zero-inflation. Sebagai variabel prediktor pada model zero-inflated ordered probit adalah jumlah kendaraan yang terlibat dan jumlah korban jiwa. Hasil penelitian ini menyimpulkan bahwa model zero-inflated ordered probit dengan jumlah kendaraan yang terlibat sebagai inflation variable dan jumlah korban jiwa sebagai variabel prediktor merupakan model terbaik. Model ini memberikan prediksi yang sesuai dengan nilai sebenarnya dengan akurasi 85,36%. Diperoleh pula nilai AIC dari model ini sebesar 2614,282.

.....Regression analysis is a statistical method used to determine the pattern of relationship between a response variable and one or more predictor variables. The regression model that is often used is the linear regression model with the assumption that the response variable is normally distributed. The development of the linear regression model is the Generalized Linear Model (GLM). One of the components of GLM is a link function which is used to connect response variables with linear predictors. The choice of this connecting function depends on the type of response variable. For categorical response variables where there is no ordering (nominal), one GLM that can be used is a model with probit link function. However, this model cannot be used for categorical response variables with ordinal order. For this reason, ordered probit

model was developed to analyze ordinal response variables with a probit link function. However, if category 0 has more observations than other categories, then zero-inflation occurs in the data response. This can cause overdispersion which results in misinterpretation of the model and wrong conclusions. For this reason, it is necessary to develop ordered probit model, namely zero-inflated ordered probit model. Zero-inflated ordered probit model parameter estimation was estimated using maximum likelihood method. The implementation of zero-inflated ordered probit model is used to predict the severity of injuries resulting from traffic accidents based on traffic accident data. As an ordinal response variable in the zero-inflated ordered probit model is the severity of injury which has three categories, namely slight injury, serious injury, and fatal injury. In this data, the category of slight injury (0) has more observations than other categories, resulting in zero-inflation. As predictor variables in the zero-inflated ordered probit model are the number of vehicles involved and the number of casualties. The results of this research conclude that the zero-inflated ordered probit model with the number of vehicles involved as the inflation variable and the number of fatalities as the predictor variable is the best model. This model provides predictions that match the actual values with an accuracy of 85,36%. The AIC value of this model was also obtained at 2614,282.