

INTISARI

MODEL REGRESI POISSON *HURDLE* UNTUK MENGATASI OVERDISPERSI AKIBAT *EXCESS ZEROS*

Oleh

Muhammad Iqbal Jinnan

12/331344/PA/14608

Regresi Poisson merupakan salah satu *Generalized Linear Model* (GLM) yang dapat memodelkan jenis data cacah. Dalam regresi Poisson mengasumsikan bahwa nilai mean dan variansi sama, biasa disebut equidispersi. Namun, dalam beberapa kasus, terdapat nilai nol yang terlalu banyak pada variabel respon. Hal tersebut menyebabkan nilai variansi tidak lagi sama dengan mean. Nilai variansi lebih besar daripada mean, biasa disebut overdispersi. Model regresi Poisson tidak lagi cocok untuk memodelkan jenis data seperti ini. Sehingga digunakan model regresi Poisson *Hurdle*. Terdapat dua bagian dalam model tersebut. Pertama, sebuah proses biner dimana terdapat dua pilihan, apakah proses membangkitkan nilai nol atau nilai positif. Bagian kedua, proses membangkitkan nilai positif saja. Kedua bagian, masing-masing akan diestimasi dengan menggunakan model biner dan *zero-truncated count*. Metode estimasi yang digunakan adalah metode maksimum likelihood.

Kata kunci : GLM, regresi Poisson, overdispersi, model *Hurdle*, model regresi Poisson *Hurdle*.

ABSTRACT

HURDLE POISSON REGRESSION MODEL FOR HANDLING OVERDISPERSION DUE TO EXCESS ZEROS

by

Muhammad Iqbal Jinnan

12/331344/PA/14608

Poisson regression is one of the Generalized Linear Model (GLM) that can model the kind of data count. Poisson regression assumes that the mean and variance are equal, or it's called equidispersion. However, in many cases, there are too many zero values on the response variable. This causes the value of variance is no longer equal to the mean. Values of variance greater than the mean, it's called overdispersion. Poisson regression model is not suitable anymore for this kind of data. Thus, it is suggested to use a Hurdle Poisson regression for handle overdispersion problem. There are two parts in the Hurdle Poisson regression model. First, a binary process where there are two options, whether the process generate a zero value or a positive value. The second part, the process generate positive value only. Two part of the model, each will be estimated by using binary model (first part) and zero-truncated count model (second part). The method of estimation use maximum likelihood method.

Keyword : GLM, Poisson regression, overdispersion, Hurdle model, Hurdle Poisson regression model.