

INTISARI

REGRESI *GENERALIZED POISSON-LINDLEY* UNTUK MENANGANI KASUS OVERDISPERSI PADA RESPON DATA CACAH (Studi Kasus: Pemodelan Faktor Risiko Jumlah Kematian Ibu per Penduduk Perempuan Usia Subur di Provinsi Jawa Timur Tahun 2021)

Angkie Hendrata
18/427690/PA/18650

Regresi Poisson adalah analisis yang umumnya digunakan dalam memodelkan variabel respon berupa data cacah. Terdapat asumsi utama yang harus dipenuhi di dalam model regresi Poisson, yaitu ekuidispersi (keadaan dimana mean dan variansi dalam model regresi Poisson bernilai sama). Pelanggaran asumsi yang umum terjadi pada regresi Poisson adalah kondisi overdispersi, khususnya pada bidang penelitian biostatistika. Kondisi overdispersi dapat mengakibatkan terjadinya bias dalam estimasi parameter dan menyebabkan model menjadi tidak valid. Salah satu penanganan kasus overdispersi pada model regresi data cacah adalah menggunakan regresi *Generalized Poisson-Lindley* yang merupakan bagian dari *Mixed Poisson*. Estimasi parameter dalam regresi *Generalized Poisson-Lindley* menggunakan metode *Maximum Likelihood* dengan iterasi BFGS.

Regresi *Generalized Poisson-Lindley* diterapkan pada studi kasus, yaitu pemodelan faktor risiko jumlah kematian ibu per penduduk perempuan usia produktif di Provinsi Jawa Timur tahun 2021. Pada masa pandemi, Jawa Timur menempati peringkat satu di dalam kasus kematian Ibu di Pulau Jawa. Diperoleh kesimpulan bahwa model regresi *Generalized Poisson-Lindley* dapat menangani overdispersi pada respon data cacah dengan nilai -2LogLikelihood , *Akaike Information Criterion*, dan *Bayesian Information Criterion* yang lebih rendah dibandingkan model regresi Poisson, serta skala dispersi berdasarkan nilai *Pearson Residuals* tidak melebihi *cut-off point*-nya. Faktor yang mempengaruhi kematian ibu adalah dampak tidak langsung dari masa pandemi, yaitu persentase persalinan ditolong tenaga kesehatan, persentase perempuan yang menikah dini (di bawah 17 tahun), dan jumlah Puskesmas.

Kata Kunci : Overdispersi, Regresi Poisson, Distribusi *Mixed Poisson*, Regresi *Generalized Poisson-Lindley*, Jumlah Kematian Ibu.

ABSTRACT

GENERALIZED POISSON-LINDLEY REGRESSION MODEL TO HANDLE THE PROBLEM OF OVERDISPERSION IN COUNT DATA RESPONSE (Case Study: Risk Factor Modelling for Maternal Mortality Total per Women in Reproductive Age in East Java Province 2021)

Angkie Hendrata
18/427690/PA/18650

Poisson regression is an analysis that is commonly used for modelling response variable in the form of count data. There is a main assumption that must be fulfilled in the Poisson regression, which is equidispersion (a condition when mean and variance in the Poisson regression is having same value). The violation of the equidispersion assumption in the Poisson regression is overdispersion, generally happens in biostatistic research field. Overdispersion condition can cause biased parameters estimation and become an invalid model. One way to handle overdispersion case in count data regression is using Generalized Poisson-Lindley regression, which comes from Mixed Poisson. The estimation of model parameters is carried out using the Maximum Likelihood method by optimizing it through BFGS iteration.

Generalized Poisson-Lindley regression is applied in a case study, that is risk factor modelling for maternal mortality total per woman in their reproductive age in East Java Province 2021. East Java Province ranked first in the maternal mortality case on Java Island during pandemic. It can be concluded that Generalized Poisson-Lindley regression model can handle overdispersion in count data response from -2LogLikelihood , Akaike Information Criterion, and Bayesian Information Criterion value that are smaller than Poisson regression, as well as dispersion scale based on Pearson Residuals does not exceed its cut-off point. Factor that influenced maternal mortality are indirect impact from pandemic, there are percentage of childbirth helped by healthcare workers, percentage of woman being married before 17, and total of public health center.

Keywords : Overdispersion, Poisson Regression, Mixed Poisson Distribution, Generalized Poisson-Lindley Regression, Maternal Mortality Total.