

Teknik imputasi ganda dengan metode generalized linear model poisson regression = Multiple imputation technique with generalized linear model poisson regression method

Muhamad Ihsan, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20494612&lokasi=lokal>

Abstrak

Nilai hilang merupakan suatu masalah yang sering dijumpai di berbagai bidang dan harus diatasi untuk memperoleh inferensi statistik yang baik seperti penaksiran parameter. Nilai hilang dapat ditemukan pada setiap jenis data, salah satunya pada jenis data cacah/ count data yang berdistribusi Poisson. Solusi untuk mengatasi masalah nilai hilang berjenis data cacah tersebut dapat diatasi dengan menerapkan teknik imputasi ganda. Teknik imputasi ganda merupakan suatu cara mengatasi nilai hilang dengan mengganti setiap nilai yang hilang dengan beberapa nilai estimasi. Teknik imputasi ganda untuk kasus data cacah terdiri dari tiga tahap utama yaitu tahap imputasi berdasarkan model linier normal, tahap analisis dengan metode generalized linear model Poisson regression dan tahap penggabungan pooling parameter yang didasarkan pada aturan Rubin. Studi ini juga dilengkapi dengan simulasi numerik yang bertujuan untuk komparasi akurasi berdasarkan nilai bias yang dihasilkan. Parameter yang digunakan pada simulasi ini yaitu sebesar 5, 10 dan 15 dengan jumlah sampel sebesar 200 untuk tujuan mengaproksimasi sifat kenormalan dan simulasi ini diulang untuk empat skenario yang bertingkat untuk setiap parameter berdasarkan besarnya persentase observasi nilai hilang (0%, 10%, 20% dan 30%). Berdasarkan studi literatur dan simulasi numerik yang dilakukan, solusi yang diajukan untuk mengatasi nilai hilang pada data cacah menghasilkan hasil yang cukup memuaskan terutama saat parameter bernilai besar dan persentase observasi nilai hilang yang kecil. Hal ini diindikasikan dengan ukuran bias dan variansi total dari taksiran rata-rata yang kecil. Namun nilai bias cenderung meningkat seiring meningkatnya persentase observasi nilai yang hilang dan saat nilai parameter yang kecil.

<hr>

Missing values are a problem that is often encountered in various fields and must be addressed to obtain good statistical inference such as parameter estimation. Missing values can be found in any type of data, included count data that has Poisson distributed. One solution to overcome that problem is applying multiple imputation techniques. The multiple imputation technique is a way of dealing with missing values by replacing each missing value with some estimated values. The multiple imputation technique for the case of count data consists of three main stages, namely the imputation stage based on the normal linear model, the analysis stage using the generalized linear model Poisson regression and the last stage is pooling parameter based on Rubins rules. This study is also equipped with numerical simulations which aim to compare accuracy based on the resulting bias value. The parameters used in this simulation are 5, 10 and 15 with a sample size of 200 for the purpose of approximating normal properties and this simulation is repeated for four multilevel scenarios for each parameter based on the percentage of observation of missing values (0%, 10%, 20% and 30%). Based on the study of literature and numerical simulations carried out, the solutions proposed to overcome the missing values in the count data yield satisfactory results, especially when the parameters are large and the percentage of observation of the missing values is small. This is indicated by the size of the bias and the total variance of the small average estimate. But the bias value tends to increase

with increasing percentage of observation of missing values and when the parameter values are small.