

INTISARI

ANALISIS REGRESI LINEAR GANDA DENGAN DATA HILANG MENGGUNAKAN IMPUTASI MULTIVARIAT CHAINED EQUATION (MICE)

Oleh:

ERLINA PUSPITARINI

07/253107/PA/11526

Data hilang sering terjadi dalam sebuah pengamatan. Seringkali peneliti membuang observasi yang memuat data hilang. Akan tetapi telah ditemukan berbagai metode untuk menangani masalah data hilang antara lain dengan imputasi. Imputasi terdiri dari imputasi klasik dan imputasi modern. Terkadang nilai imputasi sering tidak masuk akal. Seperti jumlah yang bernilai negatif, apalagi jika kasusnya merupakan data kategorik akan sangat mudah sekali terjadi kesalahan yang sangat fatal. Oleh karena itu muncul metode baru dalam imputasi yaitu menggunakan *chained equation* (MICE), karena prosedurnya berantai sehingga meminimalisir timbulnya data yang tidak masuk akal. Dalam analisis MICE setelah dilakukan iterasi sebanyak 5 kali maka akan didapat model regresi fit sebanyak 5 kali juga. Kemudian dari kelima model fit tersebut disimpulkan menjadi satu model yang disebut sebagai model pool.

Kata kunci : data hilang, imputasi, *chained equation*, regresi linear

ABSTRACT

MULTIPLE LINEAR REGRESSION ANALYSIS WITH MISSING DATA USING MULTIVARIATE IMPUTATION CHAINED EQUATION(MICE)

BY:

ERLINA PUSPITARINI

07/253107/PA/11526

Missing data often occur in an observation. Often researchers discard observations containing missing data. But has found a variety of methods for dealing with missing data, among others, by imputation. Imputation consisted of imputing classical and modern imputation. Sometimes the value of imputation often unreasonable. Such amount is negative, especially if the case is categorical data will be very easy to very fatal error occurs. Therefore, it appears that a new method of imputation using chained equations (MICE), because the procedure is a chain so as to minimize the emergence of data that does not make sense. In the analysis performed iterations MICE after 5 times it will get the regression model fit as much as 5 times as well. Later than the fifth model fit inferred from giving a model called as a model pool.

Keywords: missing data, imputation, chained equations, linear regression