

INTISARI

PENANGANAN *OUTLIER* PADA DATA MULTIVARIAT DENGAN ESTIMATOR DETMCD, DETS DAN DETMM

Oleh

FEBRIANI ASTUTI

16/403744/PPA/05261

Analisis multivariat adalah cabang dari ilmu statistika yang menggunakan data dari banyak variabel yang selanjutnya disebut data multivariat. Banyaknya variabel yang digunakan dalam penelitian berpotensi mengandung *outlier*. Adanya *outlier* berpengaruh pada estimasi nilai parameter yang bersifat bias, sehingga menyebabkan interpretasi hasil yang diperoleh menjadi tidak akurat. Akan tetapi pada saat tertentu, *outlier* tidak dapat dihapus, karena *outlier* mengandung informasi penting yang tidak dapat diberikan oleh observasi lain. Metode yang digunakan untuk memperoleh estimator terbaik yang tahan terhadap *outlier* yaitu metode *robust*. Metode *robust* yang digunakan dalam penelitian ini adalah metode *robust* dengan estimator DetMCD, DetS dan DetMM. Ketiga estimator tersebut merupakan estimator deterministik yang mampu menaikkan *kerobustan* dengan cepat pada dimensi yang tinggi. Langkah awal dari estimator deterministik adalah menghitung standarisasi setiap variabel, setelah itu menghitung enam variansi yang deterministik. Langkah selanjutnya menghitung estimasi mean dan variansi dari masing-masing variabel. Estimator DetMCD, DetS dan DetMM juga dapat mendeteksi adanya *outlier* pada data serta mengurangi efek *masking* dan *swamping* yang terjadi apabila menggunakan estimator klasik. *Masking* (penyamaran) yaitu observasi yang sebenarnya merupakan observasi *outlier* akan tersamarkan sebagai observasi yang bukan merupakan *outlier*, sedangkan *swamping* (pelimpahan) adalah observasi yang sebenarnya bukan merupakan *outlier* terdeteksi sebagai *outlier*. Dari hasil penelitian diperoleh bahwa estimator DetMCD, DetS dan DetMM mampu mendeteksi *outlier* dengan baik serta menghasilkan estimasi mean dan variansi yang lebih baik dari estimator klasik.

Kata Kunci: multivariat, *outlier*, *robust*, estimator DetMCD, estimator DetS, estimator DetMM

ABSTRACT

HANDLING OUTLIER ON MULTIVARIATE DATA WITH DETMCD, DETS AND DETMM ESTIMATORS

By

FEBRIANI ASTUTI

16/403744/PPA/05261

Multivariate analysis is a branch of statistical science that uses data from many variables, here in after called multivariate data. The number of variables used in the study has the potential to contain outliers. The existence of outliers influences the estimation of parameter values that are biased, which causes the interpretation of the results to be inaccurate. However, at certain times, outliers cannot be removed, because outliers contain important information that other observations cannot offer. The method used to get the best estimator that is resistant to outliers is the robust method. The robust method used in this study is the robust method with DetMCD, DetS and DetMM estimators. The three estimators are deterministic estimators that can increase robustness quickly in high dimensions. The first step of the deterministic estimator is to calculate the standardization of each variable, after which it calculates six deterministic variances. The next step calculates the estimated mean and variance of each variable. DetMCD, DetS and DetMM estimators can also detect outliers in the data and reduce the masking and swamping effects that occur when using a classic estimator. Masking, ie actual observation is an outlier observation will be obscured as an observation that is not an outlier, while swamping is an observation that is actually not an outlier detected as an outlier. From the results of the study it was found that the DetMCD, DetS and DetMM estimators were able to detect outliers well and produce estimates of mean and variance better than classical estimators.

Keywords: multivariate, outlier, robust, DetMCD estimator, DetS estimator, DetMM estimator