

The importance of the scales in heterogeneous robust clustering

L.A. GARCÍA-ESCUADERO* AND A. GORDALIZA

Dpto. de Estadística e Investigación Operativa, Universidad de Valladolid, Spain.

Abstract

This work is aimed to show the important role that the estimation of the scales plays in the derivation of robust cluster techniques. This holds specially for some recently proposed methods including the so-called “concentration” steps (Rousseeuw and van Driessen 1999) in their implementation.

We introduce a new robust clustering approach: the SSC method. This method starts from a high trimming level which surely serves to remove all the outlying observations. Later, an iterative process is carried out where special attention is paid to the proper estimation of the groups’ scales. The estimation of the contamination level is also considered.

Key words: Robustness; Cluster Analysis; Minimum Covariance Determinant; trimming; scale parameters.

* *Corresponding author.* Fax: +34-983-423111

E-mail addresses: lagarcia@eio.uva.es (L.A. García-Escudero), alfonsog@eio.uva.es (A. Gordaliza).