

Corsi di Laurea in Scienze Statistiche ed Attuariali

Dipartimento DEMM – Università degli Studi del Sannio

Seminario

Trimming and Wild Trimming procedures for robust clustering with mild and gross outliers

Alessio Farcomeni

Sapienza Università di Roma

Dipartimento di Sanità Pubblica e Malattie Infettive

Dipartimento DEMM, Aula ex Consiglio, piano 1, P.zza Arechi II
Mercoledì 16 Ottobre, ore 16:15

Abstract

Trimming principles play a central role in robust statistics. In this seminar we recap the main uses of impartial trimming for robust cluster analysis, with special attention to three recent developments. In the first part we introduce a trimmed mixture of contaminated Gaussian densities, which is useful for increasing efficiency by incorporating mild outliers into the estimation set. A penalized likelihood procedure with a theoretically optimal penalty parameter allows us to select the optimal trimming proportion, also in the more classical case of Gaussian mixtures (tclust). We then remove the necessity of selecting a trimming parameter in advance by using an iteratively reweighted version of the procedure, initialized from a high trimming level and then refined as possible. The resulting procedure is highly efficient despite a breakdown point corresponding to the initial trimming level. Finally, we remove the necessity of specifying both the trimming parameter and the number of clusters through an iterative clustering procedure rooted in the Forward Search approach, which replaces the original K-population (robust) estimation problem with K distinct one-population steps. The procedure is grounded on novel theoretical results concerning wild trimming, that is, trimming more than 50% of the observations. The three methods are illustrated on an original data example concerning the identification of source from illicit drug shipments seized in Italy and Spain.

Based on joint works with Andrea Cerioli, Francesco Dotto, Luis-Angel Garcia-Escudero, Agustín Mayo-Iscar, Antonio Punzo, Marco Riani.