



REPOSITORIO DIGITAL UNIVERSITARIO (RDU-UNC)

Robust clustering of banks in Argentina

Margarita Díaz, José M. Vargas, Fernando García

Ponencia presentada en XI Congreso Latinoamericano de Sociedades de Estadística. XLI Jornadas Nacionales de Estadística. XLII Coloquio Argentino de Estadística realizado en 2014 en la Universidad de La Serena. La Serena, Chile



Robust Clustering of Banks in Argentina

Margarita Díaz^{1*}, José M. Vargas^{1,2} and Fernando García
¹Department of Statistics, FCE UNC, Córdoba, Argentina,
²Department of Mathematics, IAPCB UNVM, Villa María, Argentina

Abstract

The purpose of this paper is to classify and characterize 64 banks, active as of 2010 in Argentina, by means of robust techniques used on information gathered during the period 2001-2010. Based on the strategy criteria established in [Wang (2007)] and [Werbin (2010)], seven variables were selected. In agreement with bank theory, four "natural" clusters were obtained, named "Personal", "Commercial", "Typical and "Other banks", using robust K-means clustering as implemented in R statistical language through the function [Kondo (2011)] detecting six outliers in the process. In order to characterize each group, projection pursuit based robust principal component analysis, [Croux (2005)], was conducted on each cluster revealing approximately a similar component structure explained by three components in excess of 80 %, granting a common principal components analysis as in [Boente (2002)]. This allowed us to identify three variables which suffice for grouping and characterizing each cluster. Boente influence measures were used to detect extreme cases in the common principal components analysis.

Keywords: Robust clustering Projection pursuit Common Principal Components Influence measures.

^{*}Email: pdiaz@fcm.unc.edu.ar, Corresponding author