

Nuevos enfoques al problema de encontrar grupos, atípicos y otras formas de heterogeneidad en una muestra

por

Daniel Peña

Universidad Carlos III de Madrid

Los problemas de encontrar atípicos y detectar grupos en una muestra se han tratado tradicionalmente en la literatura por procedimientos distintos. Los atípicos se han estudiado por procedimientos diagnósticos y métodos robustos, mientras que encontrar grupos en una muestra de datos multivariantes es un tema central en la literatura de cluster analysis. En los últimos años varios trabajos han emprendido la tarea de unificar ambos problemas bajo el paraguas común de la detección de la heterogeneidad en una muestra. Este charla presenta algunos de estos enfoques desarrollados en los últimos cinco años. En particular se presentan métodos basados en proyecciones que buscan los extremos del coeficiente de curtosis de la proyección, tipo projection pursuit, para identificar atípicos y encontrar clusters, y métodos de construir grupos basados en ideas de validación cruzada donde la muestra se divide en grupos que después se recombinan para identificar la heterogeneidad. Estos métodos sirven también para detectar atípicos.

La charla esta basada en numerosos ejemplos de aplicación, en Economía, Ingeniería y Biología.