

Fundamentos del análisis de componentes principales

El ACP encuentra un nuevo conjunto de variables, denominados “componentes principales”, no correlacionadas entre sí (ortogonales), que explican la estructura de variación de los casos observados con una pérdida mínima de información (Balzarini et al., 2008). El sistema de coordenadas que explica la mayor varianza del conjunto de datos se denomina componente 1, mientras que la segunda coordenada, componente 2, y así sucesivamente. La proporción de variabilidad total explicada por estas componentes (o “lambdas”) se expresa como porcentaje, logrando reducir las dimensiones de los datos a dos o más ejes cartesianos. En los gráficos resultantes (“*bi-plot*”), las variables se grafican como vectores desde el origen (intersección entre componentes) y las observaciones, como puntos (manuales escolares). Conviene aclarar que la distancia entre las observaciones y las variables no tiene interpretación en el ACP, mientras que la dirección de éstas al origen sí aporta valiosa información. Así, los ángulos agudos y obtusos entre los vectores pueden ser interpretados como correlaciones entre variables (positivas y negativas, respectivamente), mientras que ángulos de 90° indican que las variables son independientes. Tanto en el caso argentino como español, las subcategorías de la Tabla 2 fueron consideradas “variables” y los códigos de identificación de los manuales, “criterio de clasificación” (opcional para el ACP). Para realizar el ACP utilicé el programa Infostat (versión 2013).