

FACULTAD DE INGENIERÍA

ESTIMADORES DE POSICIÓN Y ESCALA MULTIVARIADOS DE TIPO MM EN PRESENCIA DE DATOS FALTANTES

Marfia, Martín

Alvarez, Enrique (Dir.)

Facultad de Ingeniería, UNLP.

martinmarfia@hotmail.com

PALABRAS CLAVE: Datos Faltantes, Estimador Robusto, Alta Eficiencia.

MM TYPE LOCATION AND SCALE MULTIVARIATE ESTIMATORS IN PRESENCE OF MISSING DATA

KEYWORDS: Missing Values, Robust Estimator, High Efficiency.

Resumen gráfico



Resumen

La mayoría de los procedimientos estadísticos clásicos están basados en modelos con hipótesis rígidas, tales como errores normales, observaciones equidistribuidas, etc. Bajo estas hipótesis se deducen procedimientos óptimos. Por ejemplo, para el caso de regresión el procedimiento óptimo es el de mínimos cuadrados; para modelos paramétricos en general, los procedimientos óptimos clásicos son los estimadores de máxima verosimilitud. Sin embargo, estos métodos son muy sensibles al incumplimiento de las hipótesis que los generaron, tales como la presencia en la muestra de observaciones atípicas. En efecto, pequeñas desviaciones a la normalidad manifestadas por unas pocas desviaciones atípicas pueden invalidar las conclusiones basadas en estos procedimientos. Los procedimientos estadísticos robustos tienen como objetivo permitir inferencias válidas cuando el modelo no se cumple exactamente y al mismo tiempo ser altamente eficientes bajo el modelo. A la falta de robustez de los métodos de estimación clásicos se le suma otro problema común como son los datos faltantes. La presencia de datos faltantes, es la situación a la que permanentemente se enfrentan investigadores, ingenieros, analistas, etc. El objetivo es disponer de un

archivo de datos completos, pero aplicar métodos de imputación inapropiados para lograrlo, puede generar más problemas de los que resuelve. La solución más simple es restringir la atención a los casos completos, pero esto no haría un uso eficiente de los datos y en el caso de muestras pequeñas no resulta viable. Durante las últimas décadas se han desarrollado procedimientos que tienen mejores propiedades estadísticas que las opciones tradicionales.

Asumimos que nuestros datos están completamente perdidos al azar (MCAR), lo que quiere decir que las componentes de una observación se pierdan no depende de los valores que tome esa observación. Nuestro objetivo es proponer una generalización del estimador de tipo MM para el modelo de posición y escala multivariado que sea capaz de enfrentar ambos problemas. Para ello, nuestro enfoque es considerar como estimador de escala inicial la escala de las distancias de Mahalanobis parciales del S-estimador generalizado, y usarlo como punto de partida para calcular un M-estimador cuya función rho tiene un parámetro para controlar la eficiencia.

Multimedia

<http://sedici.unlp.edu.ar/handle/10915/114055>