

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).

martin.marfia@ing.unlp.edu.ar

PALABRAS CLAVE: robustez, datos faltantes, datos multivariados.

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

KEYWORDS: robustness, missing values, multivariate data.

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.

A pesar de que la presencia de valores atípicos y datos faltantes han sido individualmente muy estudiados, son pocos los trabajos que estudian ambos problemas en conjunto.

Recientemente, Danilov y otros (2012) propusieron dos estimadores robustos para datos faltantes: un S-estimador generalizado (GSE) y un S-estimador extendido (ESE). Ambos estimadores coinciden con los S-estimadores propuestos por Davies (1987) para muestras sin datos faltantes.

Nuestra propuesta consiste en un estimador de posición y escala donde en la primera etapa utilizamos el estimador propuesto por Danilov. Nuestro objetivo es estudiar condiciones de existencia, consistencia, punto de ruptura y normalidad asintótica. Además, queremos proponer un algoritmo para el cálculo numérico del estimador para poder realizar simulaciones y poder compararlo con otros ya existentes.