

## Comunicaciones

**Título:** Robust clustering methods based on the tau scale: preliminary results.

**Autores:** Juan Gonzalez, Victor Yohai and Ruben Zamar

**Resumen:** Cluster Analysis methods aim to detect homogeneous groups within a population. It is well known that a single outlier in the data may have a large effect on the selected clusters when they are obtained using a traditional procedure as, for example, the K-Means algorithm. To overcome this problem, we propose a clustering procedure based on the minimization of an efficient and robust scale (the tau scale proposed by Yohai and Zamar (1988)) of the distances between the points and the corresponding cluster center.

To evaluate the proposed clustering algorithm, numerical experiments have been carried using several data sets. The results show that the proposed procedure based on the minimization of the tau-scale compares favorably with respect to other widely used algorithms as K-Means and Trimmed-K-Means.

**Título:** A robust regression estimator with high finite sample efficiency.

**Autores:** Ezequiel Smucler – Victor Yohai

**Resumen:** MM-estimators, introduced in Yohai (1987), are regression estimators that can be tuned to attain both a high breakdown point and an asymptotic efficiency at the normal distribution as close to one as desired. However, as will be seen in the simulation study we report, when the sample size is not very large, MM-estimators can have a finite sample efficiency that is much lower than their asymptotic one.

We propose a new estimator, and prove its consistency and maximum attainable finite-sample replacement breakdown point. Results of a simulation study show that compared to MM-estimators, the new estimator has relatively high efficiency for small sample sizes and a comparable behavior in the presence of contaminated data.

**Título:** M-estimadores basados en transformaciones.

**Autores:** Marina Valdora

**Resumen:** En este trabajo se estudia una familia de estimadores para los modelos lineales generalizados que tienen una definición simple y son altamente robustos. Los estimadores propuestos son M-estimadores redescendientes aplicados a las respuestas transformadas (MT-estimadores). El objetivo de transformar las respuestas es estabilizar sus varianzas en un valor casi constante y así permitir un correcto escalamiento de la función de pérdida utilizada para definir el M-estimador.

**Título:** a confirmar.

**Autores:** Luis Foglia

## Posters

**Título:** Un Test de Simetría en Espacios de Dimensión Infinita.

**Autores:** Fraiman Ricardo, Moreno Leonardo, Vallejo Sebastian.

**Resumen:** Se presenta un test de simetría central para datos multivariados mediante proyecciones al azar, basado en un test univariado del tipo Kolmogorov Smirnov. Se prueba la distribución libre del test bajo la hipótesis nula y su consistencia bajo cualquier alternativa. Se proponen diferentes mecanismos para mejorar la potencia de la prueba. Mediante un estudio de simulación se evalúa su eficiencia respecto a otros test de simetría multivariados. Se realizan extensiones a espacios de Hilbert y de Banach.

**Título:** Inclusión de ASC en modelos econométricos para medir preferencias de consumo

**Autores:** Beatriz Lupín & Elsa M. Rodríguez

Grupo de Investigación "Economía Agraria"

Facultad de Ciencias Económicas y Sociales–Universidad Nacional de Mar del Plata  
beatrizlupin@gmail.com / emrodri@mdp.edu.ar

**Resumen:** La metodología *Choice Modelling* enfrenta a los consumidores a alternativas de un bien–conformadas por combinaciones de atributos del mismo– y a la alternativa “no compraría ninguna de las opciones anteriores” (*opt-out*). Luego, los participantes deben elegir alguna/s de las alternativas posibles.

A partir de las preferencias declaradas por los individuos, es posible especificar modelos de elección discreta para estimar la utilidad derivada del consumo de un producto, asumiendo que la misma es aleatoria.

Existen diversas posturas acerca de la inclusión de Constantes Específicas por Alternativa (ASC), las que representan fuentes sistemáticas de utilidad no observables en las elecciones de los participantes.

Hensher *et al.* (2007) sostienen que, en el caso de *unlabeled experiments*<sup>1</sup>, no es necesario introducirlas porque carece de sentido su interpretación. Sin embargo, otros autores indican que al excluirlas, el resto de los parámetros captarán el efecto de los factores inobservables, resultando sesgados (Hoyos, 2010). Particularmente, lo anterior se evidencia en presencia de la *opt-out*, por lo que recomiendan añadir una ASC asociada a ella (Kontoleon & Yabe, 2003; Scarpa *et al.*, 2005).

En este trabajo, se propone una revisión teórico–conceptual y una reflexión crítica acerca de la inclusión de ASC en Modelos de Regresión Logística.

Se trata de experimentos *Choice Modelling* en los que las alternativas del producto en cuestión se presentan a los participantes sin un rótulo específico a fin de evitar cualquier tipo de influencia en la elección. Así, por ejemplo, se denomina cada alternativa con una letra.

**Título:** Construcción de funciones de estimación múltiple protegidas en modelos con verosimilitud factorizada.

**Autor:** Julieta Molina –UBA

**Resumen :** Recientemente desarrollamos un marco teórico que explica la existencia de estimadores múltiple protegidos en varias de las propuestas de estimación múltiple protegidas ya existentes, en modelos donde la verosimilitud se factoriza. Dicha teoría provee condiciones suficientes bajo las cuales es posible

---

<sup>1</sup>

la construcción de ecuaciones de estimación que, bajo condiciones de regularidad, tienen soluciones que constituyen estimadores múltiple protegidos. Presentamos también condiciones suficientes bajo las cuales es posible estimar de manera múltiple protegida ciertos parámetros de ruido, cuya estimación es necesaria para poder estimar de manera múltiple protegida al parámetro de interés. En este poster damos algunos lineamientos acerca de como arribamos a estos procedimientos de estimación. Además presentamos estimadores doble y múltiple protegidos para la componente lineal bajo un modelo de regresión parcialmente lineal con datos faltantes.

**Título:** Robust Tests in the Nonlinear Model for Complete and Incomplete Data

**Autor:** Paula Mercedes Spano – CONICET – UBA

**Resumen :** Consideramos un modelo de regresión no lineal en el que puede haber respuestas faltantes, que pueden ocurrir simplemente por diseño o por azar. En este contexto, los métodos de inferencia clásicos sobre el parámetro de regresión se basan en mínimos cuadrados, sin embargo es bien sabido que resultan muy vulnerables a la presencia de datos atípicos.

Proponemos una familia de M-estimadores pesados, que resultan consistentes y asintóticamente normales, para estimar el parámetro de regresión de un modelo no lineal en el que puede haber respuestas faltantes al azar.

Asimismo, a fin de realizar pruebas de hipótesis que involucran al parámetro de regresión, introducimos una familia de tests robustos utilizando estadísticos de tipo Wald basados en estos M-estimadores pesados. Estudiamos la distribución asintótica del estadístico propuesto bajo la hipótesis nula y bajo alternativas contiguas.

Llevamos a cabo un estudio de Monte Carlo para evaluar el comportamiento en muestras finitas de los procedimientos presentados en datos contaminados y sin contaminar bajo diferentes modelos de respuestas faltantes

**Título:** Estimacion de Parametros en Modelos con Errores en las Variables

**Autor:** Manuel Benjamin – UBA–INTI

**Título:** Searching for the core variables in principal components analysis Autor:

**Autor:** Guido Guissani – CONICET – UdeSA

**Resumen :** We introduce a procedure for selecting variables in principal components analysis. The procedure was developed to identify a small subset of the original variables that "best explain" the principal components through nonparametric relationships. There are usually some noisy uninformative variables in a dataset, and some variables that are strongly related to each other because of their general interdependence.

The procedure is designed to be used following the satisfactory initial use of a principal components analysis with all variables, and its aim is to help to interpret underlying structures. We analyze the asymptotic behavior of the method and provide two examples.

**Título:** Seeking Relevant Information from a Statistical Model

**Autor:** Yanina Gimenez – CONICET–UdeSA

**Resumen:** We herein introduce a general procedure for selecting variables, which can be applied to several classical multivariate problems, including principal components, regression, classification, clustering, among others.

The aim is to allow the identification of a small subset of the original variables that can "better explain" the model concerned through nonparametric relationships. The method typically yields some "noisy" uninformative variables and some variables that are strongly related because of their general dependence and our aim is to help understand the underlying structures in a given dataset. The asymptotic behaviour of the proposed method is considered and some data are analysed as examples.