Department of Decision Sciences - Bocconi University
Via Roentgen 1 - 20136 Milano
Tel. 02-58365632 - Fax 02-58365630

SEMINAR

"A multivariate data mining approach to measure selection bias and test balance"

Furio Camillo
(Department of Statistical Sciences, Bologna University)

Thursday, 5th March 2009 – h. 16.30
Room 35 – Via Sarfatti 25 - 20136 Milano

Abstract:

We propose a creative and practical approach to deal with the problem of selection bias, that may be the most important vexing problem in program evaluation or in any line of research that attempts to assert causality. We propose a multivariate approach that involves measuring selection bias under non-experimental conditions and performing a multivariate test of imbalance in order to check if the detected bias is significant, by simultaneously considering all covariates involved in the selection process and, thus, preserving the multivariate nature of data.
The basic idea involves investigating the dependence relationship
between a set of observable pre-treatment covariates X and a treatment indicator variable T in order to obtain a measure of selection bias according to their dependence structure.
The approach is non parametric and involves the construction of a conditional multidimensional space of the X-matrix in which the bias associated with treatment assignment has been eliminated.
The measure of selection bias is then represented in terms of variability of the original X-space that has been eliminated.
Finally, we propose the use of a clustering procedure as a tool to find groups of comparable units on which estimate local causal effects and the use of the multivariate test of imbalance as a stopping rule in choosing the best cluster solution set.

(joint work with Ida D'Attoma)