

# Seminario

## **DOTTORATO DI METODI STATISTICI PER L'ECONOMIA E L'IMPRESA (DMSEI)**

Marco Riani

Dipartimento di Economia  
Università di Parma

Titolo del seminario

### **Robust and Efficient Methods for the Analysis of Complex Data**

#### ABSTRACT

Data are an overwhelming feature of modern life. As the amount of data increases so do the challenges facing the statistician in trying to extract information from ever larger data sets. One of the arguments of the talk is that larger data sets are also more complex and require flexible multiple analyses in order to reveal their structure. Only then can information be efficiently extracted. The analysis of large data sets may be complicated by the high dimensionality of responses, large numbers of observations and complexity of the choices to be made among explanatory variables. Although appreciable, these challenges to the statistician are not different in kind from those faced in the analysis of smaller data sets. In my talk I will focus on problems that become severe in large complex data sets, such as our inability to find a single model for all the data. The simplest situation is that of a single model with possibly many outliers. In the presence of a core population and some isolated or clustered outliers, traditional robust methods can be successfully used to find proper models. However, when there are several populations and different sources of heterogeneity, traditional robust methods fail to recover the real structure of the data and more sophisticated procedures, such as those derived from the Forward Search are required.

~ ~ ~

**Martedì 9 febbraio 2010 aula 20 (II piano) - ore 11,00**

**Facoltà di Economia**  
**Università degli Studi Roma Tre**  
*Via Silvio D'Amico 77, 00145 Roma*