

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

## ***SEMINAR***

# **"Robust Random Effects Models: a diagnostic approach based on the forward search"**

**Roberta Varriale**

(Department of Statistics "G.Parenti", University of  
Florence, Italy)

**Thursday, 9th July 2009 - h. 16.30**  
**Room C - Via Sarfatti 25 - 20136 Milano**

---

### **Abstract:**

This work presents a simple robust method for the detection of atypical observations and the analysis of their effect on model inference in the random effect linear models (McCulloch and Searle, 2001). In particular, we extend the approach used in the fixed effect framework (Bertaccini and Varriale, 2007) using a Forward Search procedure that orders the observations by their closeness to the hypothesized model (Atkinson and Riani, 2000).

The starting point of the FS is to fit the model to very few observations chosen with a robust procedure, order all the observations by their closeness to the fitted model, increase the subset size and refit the model. The process continues with increasing subset sizes until all data are fitted. In random effect models, also known as multilevel models, the outliers may affect the data at each level of observation. Attention is limited to two hierarchical levels. During the search, at each stage, we monitor some informative quantities, such as parameter estimates, residual plots and other relevant statistics in order to identify the outliers. In particular, we focus on the effect of outliers on the second-level variance using the likelihood ratio test suggested by Self and Liang (1987). A cut-off point separating the outliers from the other observations is identified through a graphical analysis of the information collected at each step of the Forward Search; the Robust Forward LRT is the value of the classical LRT statistic at the cut-off point. Through some Montecarlo simulation studies we are able to claim the clear superiority of our proposal since the probability of the type I error computed with the FS method is much lower than the one computed with the classical approach when data are contaminated, without any loss in terms of power when data are not contaminated.

Work in collaboration with:

B. Bertaccini

Department of Statistics "G.Parenti", University of Florence, Italy