

Centro de Investigação em Matemática e Aplicações Mestrado em Modelação Estatística e Análises de Dados Departamento de Matemática

# <u>Seminário</u>

#### 3 de Novembro de 2011, Quinta-feira CLAV – Anf. 1 - 14:00 horas

## Analysis of cluster correlated data from complex surveys

### Georgiy Bobashev

RTI International, Durham, NC USA.

### **Resumo**

Traditional statistical inference is based on the fundamental concept of independency and identical distribution of observations. When clustering in considered in sampling or randomized clinical trials the observations are not independent anymore and a modification of more traditional methods is needed. The definition of clusters is very broad and covers primary sampling units such as neighborhoods, schools, hospitals, as well as individuals when multiple measures are taken for the same individual or entity.

In the seminar I will discuss the concept of cluster correlation and its implications in survey and clinical studies. I will discuss two major models that account for clustering: population averaged and cluster-specific (hierarchical) models. I will illustrate the basics of the generalized estimating equation (GEE) methods and the use of Alternating Logistic Regression (ALR) as an efficient way to estimate the amount of clustering of binary data.