SEMINAR

"Bayesian Methods for Genomics"

Marina Vannucci
(Department of Statistics, Rice University, Houston)

Thursday, 12th June 2008 – h. 16.30
Room 44 – Via Sarfatti 25 - 20136 Milano

Abstract:

The analysis of the high-dimensional genomics data generated by modern technologies, such as DNA microarrays, poses challenges to standard statistical methods. In this talk I will describe how Bayesian methodologies can be successfully employed in the analysis of such data. In particular, I will look at ways to incorporate "spike-and-slab" priors into various kinds of modelling settings, including linear models and model-based clustering, for the selection of the important variables (genes). Models and algorithms are quite flexible and allow the incorporation of additional information, such as data substructure. In the last part of the talk I will show how information on known biological gene functions (pathways) can be incorporated into the linear modelling frameworks for gene expression data. I will present results on simulated data and on microarray data for breast cancer prediction.