STATISTICS COLLOQUIUM

MONDAY, OCTOBER 19, 2009
TALK: 4:00 PM — SCIENCE CENTER, ROOM 309
RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

Mixed Membership Models: Tradeoffs between model complexity and Interpretability of 'shared membership'

Burton Singer
Princeton University
and
Emerging Pathogens Institute, University of Florida

ABSTRACT

Mixed membership models are ideally suited for characterizing heterogeneous populations where many individuals have multiple characteristics of interest, no combination of which occurs at high frequency in an overall population. Examples of this phenomena are: (i) representation of the health status of elderly populations at the community level; (ii) prevalence of, and variation in, co-infection with tropical diseases at the district level; (iii) characterization of environmental and behavioral risks for Chagas disease in Latin America; (iv) representation of malaria risk in complex eco-systems (e.g. the Amazon frontier, Brazil); and (v) representing variation in disability among the elderly in the United States. Although grade of membership (GoM) models have been fit to data from the above and other settings, very little attention has been given to the severe lack of interpretability of 'shared membership' of characteristics of individuals among a family of extreme, or ideal, types. To surmount this problem, we demonstrate alternative specifications of mixed membership models where an increase in dimensionality of grade of membership scores is traded for simplicity in the number and structure of ideal types, with clear payoff in terms of interpretability of model output. The overall strategy has application well outside of the epidemiological settings that we will emphasize.