Health Effects of Chemical Composition and Size of Particulate Matter: New Statistical Challenges

Francesca Dominici

Optimal regulatory control of particulate matter (PM) is hindered by an insufficient understanding of the chemical characteristics of the ambient PM mixture that determine its toxicity. Ambient PM varies spatially and temporally in its characteristics, reflecting the blend of contributing sources and meteorological factors. With regard to size, $PM_{2.5}$ and $PM_{10\text{-}2.5}$ are distinct entities with fundamentally different sources and formation processes, chemical composition, and transport distances. With respect to chemical composition, approximately 50 correlated chemical components of $PM_{2.5}$ are monitored almost daily for several hundred monitoring stations in the US. A critical question is whether a small subset of these chemical constituents, alone or through their interactions are responsible for the estimated harmful effects of PM_{10} and $PM_{2.5.}$

To estimate the toxicity of the PM complex mixture and constituents we have assembled a national data set on health outcomes, PM chemical constituents, and all the important confounders. They include: 1) the National Medicare Cohort comprising individual-level data on diseases, age, gender, and race for virtually the entire population of US elderly; 2) concentrations of key air pollutants (gaseous pollutants, PM_{10}, $PM_{2.5}$, $PM_{10\text{-}2.5}$, and $PM_{2.5}$ chemical species) ; 4) weather for all monitoring stations across the country; and 5) US Census data. These data were collected for a variety of purposes, and are misaligned in time and space.

In this talk we will present Bayesian hierarchical regression models for analyses of these large and complex national datasets. Our modeling approaches will have the ultimate goal of estimating the health effects of PM size and chemical composition on a national scale accounting for several sources of uncertainty.