



Department of Statistics Seminar

Tuesday, May 16, 2017

3:45 – 4:45 p.m., Room 420, Olmsted Hall

Reception in Olmsted 1331 at 3:15 p.m.

The Energy of Data



Gabor J. Szekely, Ph.D.

Program Director of Statistics, National Science Foundation

Research Fellow, Hungarian Academy of Sciences

Abstract:

The energy of data (E) is the value of a function of distances/dissimilarities between data. The name energy derives from Newton's gravitational potential energy which is also a function of distances between physical objects. One of the advantages of working with energy functions is that even if the observations/data are complex objects, like graphs or patterns, we can use their real valued distances for inference. The direct connection between energy and mind/observations/data is a counterpart of Einstein's $E=mc^2$. Quantifying knowledge i.e. mapping the mind (via computational neuroscience or via psychometrics) is the road to immortality (kind of: knowledge can be mapped, soul cannot). We introduce a metric (a distance) D between data such that for the energy (E) in our mind $E = FD^2$ where F is a force. This can be applied for testing equality of distributions, clustering and measuring dependence.

Biography:

Székely received his Ph. D. under the direction of Paul Erdős and Andrey Kolmogorov. Between 1990 and 1997 Székely was the founding chair of the Department of Stochastics of the Budapest Institute of Technology (Technical University of Budapest). In 1989 he was visiting professor at Yale University, and in 1990-91 he was the first Lukacs Distinguished Professor in Ohio. Székely was academic advisor of Morgan Stanley, NY, and Bunge, Chicago, helped to establish the Morgan Stanley Mathematical Modeling Centre in Budapest (2005) and the Bunge Mathematical Institute (BMI) in Warsaw (2006) to provide quantitative analysis to support the firms' global business. Since 2006 he is a Program Director of Statistics of the National Science Foundation. Székely is also Research Fellow of the Hungarian Academy of Sciences and the author of two monographs, *Paradoxes of Probability Theory and Mathematical Statistics* and *Algebraic Probability Theory*.